

Business Economics

Business Economics

**Mr Anurag Hazarika
Dr. Pradip Mahanta**

BALAJI PUBLISHERS

MEERUT-250001

Business Economics

© Author

First Edition 2022

ISBN 978-93-85756-63-4

All rights reserved. No part of this publication may be reproduced, stored in retrieval system, or transmitted, in any form or by any means, electronics, mechanical, photocopying, recording or otherwise, without the prior written permission of the author and Publisher.

Published by:

BALAJI PUBLISHERS

Head Office::

39G, 2BHK, Godwin City, Near Godwin Hotel
Near Baghpat Bypass, Meerut-250001 (U.P.)

Branch Office::

23/623, Gali No. 6, Gurana Road
Pathankot, Baraut (Bagpat)
U.P.-250611

Ph.: 09811864351, 08433295480
E-mail: balajibooks2010@gmail.com
balaji.book2010@rediffmail.com

Printed at : **India**

Preface

Business Economics is concerned with economic issue and problems related to business organization, management and strategy. Issues and problems such as the following: an explanation of why firms emerge and exist; why they expand: horizontally, vertically and spacially; the role of entrepreneurs and entrepreneurship; the significance of organizational structure; the relationship of firms with employees, the employees, the providers of capital, the customers, the government; the interactions between firms and the business environment.

If an increase in the price of a commodity causes households to expect the price of a commodity to increase further, they may start purchasing a greater amount of the commodity even at the presently increased price. Similarly, if the household expects the price of the commodity to decrease, it may postpone its purchases. Thus, some argue that the law of demand is violated in such cases. In this case, the demand curve does not slope down from left to right; instead it presents a backward slope from the top right to down left. This curve is known as an exceptional demand curve. prestigious goods also fail law of demand.

Perfect competition is an abstract concept that occurs in economics textbooks, but not in the real world. That's because it's impossible to attain in real life. Imperfect competition occurs in a market when one of the conditions in a perfectly competitive market are left unmet. This type of market is very common. In fact, every industry has some type of imperfect competition. The business cycle describes the rise and fall in production output of goods and services in an economy. Business cycles are generally measured using the rise and fall in the real gross domestic product (GDP) or the GDP adjusted for inflation.

(vi)

Demand Forecasting is the process in which historical sales data is used to develop an estimate of an expected forecast of customer demand. To businesses, Demand Forecasting provides an estimate of the amount of goods and services that its customers will purchase in the foreseeable future. Critical business assumptions like turnover, profit margins, cash flow, capital expenditure, risk assessment and mitigation plans, capacity planning, etc. are dependent on Demand Forecasting.

In a competitive market economy, price is determined by free play of demand and supply. The price will move forward or backward with changing supply and demand conditions. The going market price acts as basis for fixing the sale price. Rarely an individual seller can dishonour the current market price. In a free market economy, we have freedom of contract, freedom of enterprise, free competition and right to private property.

The book will immensely help students, researchers, teachers, policymakers, government officials and all other concerned with this subject.

— *Editor*

Contents

<i>Preface</i>	(v)
1. Introduction	1
• Basic Problems of Economy	
• Meaning, Nature and Scope of Economics	
• Characteristics of Business Economics	
• National Income and its Importance in Business Economics	
2. Demand Analysis	72
• Determinants of Demand	
• Concept of Demand Estimation	
• Demand Forecasting	
• Elasticity of Demand: Meaning Measurement, Determining Factors and Importance	
3. Production Function	124
• Meaning and Types of Production Function	
• Laws of Variable Proportions	
• Laws of Returns to Scale	
• Iso-Quants: Meaning and Properties	
• Least Cost Combination	
4. Cost Analysis	182
• Cost Concept	
• Short Run and Long run cost curves	
• Cost: Output relation, Modern Approach of Cost	
5. Market Structure and Price Determination	213
• Perfect Competition	
• Price and Output Determination	
• Importance of Perfect Competition	
• Monopoly: Discrimination Monopoly, Dumping, Comparison Between Perfect Competition and Monopoly	

(viii)

- Monopolistic Competition
- Price Determination
- Excess Capacity
- Selling Costs
- Comparison with Perfect Competition and Monopoly
- Oligopoly: Characteristics, Price Leadership, Kinked Demand Curve

6. Factor Pricing **270**

- Need of separate Theory of factor Pricing
- Marginal Productivity theory of factor Pricing
- Concepts: Rent, Wage, Interest and Profit
- Theories of Profit and Profit Policies

Bibliography **297**

Index **299**

Introduction

BASIC PROBLEMS OF ECONOMY

Basic Economic Problems

In brief, the basic economic problems may be grouped under four headings:

- The rate of output as a whole or the total national income.
- The composition and prices of output, that is, types of goods and services and their prices.
- The allocation of factors of production among their alternative uses, and
- The distribution of income as rent, profit, interest and wages.

Economics and Prices

Taking up the first, it can be suggested that economics is not entirely a science of prices. For, the area covered by prices does not embrace all the economic activities which are directed towards the satisfaction of wants. Nor does it include all the items which economists normally study. Married women do work at home. Whether they work at home or work outside for money payments, their work results into goods and services which are economic. Similarly, in a country like India most of the exchanges of goods and services take place without money and yet, these activities do lead to wealth and economic prosperity.

Problems Faced by Small Businesses

Small businesses often face a variety of problems related to their size. A frequent cause of bankruptcy is undercapitalization. This is often a result of poor planning rather than economic conditions-it is common rule of thumb that the entrepreneur should have access to a sum of money at least equal to the projected revenue for the first year of business in addition to his anticipated expenses.

For example, if the prospective owner thinks that he will generate \$100,000 in revenues in the first year with \$150,000 in start-up expenses, then he should have no less than \$250,000 available. Failure to provide this level of funding for the company could leave the owner liable for all of the company's debt should he end up in bankruptcy court, under the theory of undercapitalization.

In addition to ensuring that the business has enough capital, the small business owner must also be mindful of contribution margin (sales minus variable costs). To break even, the business must be able to reach a level of sales where the contribution margin equals fixed costs.

When they first start out, many small business owners underprice their products to a point where even at their maximum capacity, it would be impossible to break even. Cost controls or price increases often resolve this problem. In the United States, some of the largest concerns of small business owners are insurance costs (such as liability and health), rising energy costs and taxes. In the United Kingdom and Australia, small business owners tend to be more concerned with excessive governmental red tape.

Another problem for many small businesses is termed the 'Entrepreneurial Myth' or E-Myth. The mythic assumption is that an expert in a given technical field will also be expert at running that kind of business. Additional business management skills are needed to keep a business running smoothly.

Problems with using internal rate of return

As an investment decision tool, the calculated IRR should *not* be used to rate mutually exclusive projects, but only

to decide whether a single project is worth investing in. In cases where one project has a higher initial investment than a second mutually exclusive project, the first project may have a lower IRR (expected return), but a higher NPV (increase in shareholders' wealth) and should thus be accepted over the second project (assuming no capital constraints). IRR assumes reinvestment of interim cash flows in projects with equal rates of return (the reinvestment can be the same project or a different project). Therefore, IRR overstates the annual equivalent rate of return for a project whose interim cash flows are reinvested at a rate lower than the calculated IRR. This presents a problem, especially for high IRR projects, since there is frequently not another project available in the interim that can earn the same rate of return as the first project. When the calculated IRR is higher than the true reinvestment rate for interim cash flows, the measure will overestimate — sometimes very significantly — the annual equivalent return from the project. The formula assumes that the company has additional projects, with equally attractive prospects, in which to invest the interim cash flows.

This makes IRR a suitable (and popular) choice for analyzing venture capital and other private equity investments, as these strategies usually require several cash investments throughout the project, but only see one cash outflow at the end of the project (e.g., via IPO or M&A). Since IRR does not consider cost of capital, it should not be used to compare projects of different duration. Modified Internal Rate of Return (MIRR) does consider cost of capital and provides a better indication of a project's efficiency in contributing to the firm's discounted cash flow.

In the case of positive cash flows followed by negative ones and then by positive ones (for example, + " " " +) the IRR may have multiple values. In this case a discount rate may be used for the borrowing cash flow and the IRR calculated for the investment cash flow. This applies for example when a customer makes a deposit before a specific machine is built. In a series of cash flows like ("10, 21, "11), one initially invests money, so a high rate of return is best, but then receives more than one possesses, so then one owes money, so now a low rate of return is best. In this case

it is not even clear whether a high or a low IRR is better. There may even be multiple IRRs for a single project, like in the example 0% as well as 10%. Examples of this type of project are strip mines and nuclear power plants, where there is usually a large cash outflow at the end of the project.

Types of Problems and Managerial Decision Making

There are different type of problems faced during the process and progress, so a variety of decisions are implemented in accordance with situation of problem, either problem is well structured or roughly Poorly structured.

1. Well structured problems are predefined problems and are easy to tackle from the past history or practice. They can be handled by the managers through programmed decisions. Three plans are there to take programmed decisions.
 - Procedure: It is a process of continuous sequences, which can be applied in structured problem.
 - Rule: It is the guideline to the managers that whether they can do or cannot do anything.
 - Policy: It is a set of parameters that are used in making a decision includes the boundaries and limits along with restrictions.
2. Unusual or new problems with limited or insufficient information are poorly structured problems. Non programmed decisions are used to tackle such type of problems. The decision in such type of problem should be customized and unique.

General Organizational Situations

Poorly structured problems are managed through non programmed decisions by higher levels of managers within the organization and well-structured problems are solved through programmed decisions by lower levels of managers within an organization. The decisions are made on the basis of rationality, perception and bounded rationality by the managers, so we have to understand them gradually:

Assumptions of Rationality

Decision Making by the managers is assumed to be rational due to the choices that are reliable and valuable under the specified parameters. They are summarized as:

- The problems are very clear, target oriented, all alternative options are well known, clarity of preferences, parameters are non cost and time consuming and maximum profitable in these assumptions.
- Economic interests of the organization that are kept in mind in these assumptions are not of the managerial interests.

These assumptions can be achieved if; goals are clear, alternatives are limited and simple problems are faced by the manager, in which output is solid and countable and risk & innovation is supported by organizational behavior. The fault of rational model is that it's not applicable to actual decision targets for two reasons;

- Unavailability of complete information.
- Decisions are interfered by the personality factors and norms of the manager.

Actual decision making models are calculated with the presentation of the ideal rational model.

Decision-Making Styles

All the decisions are made to solve the problems by the managers with different styles and perspectives. The approach of decision making differs in two dimensions. One is individual's thoughts, i.e. rational or instinctive, and other is the (high or low) tolerance of uncertainty.

Four decision-making styles are resulted by going through two dimensions.

- The style that characterizes by low tolerance for uncertainty along with a rational way of thoughts is Directive Style.
- The style, which is characterized by a high level of tolerance for uncertainty with rational thinking way, is Analytic Style.

- Instinctive thinking way and high tolerance towards uncertainty are the characteristics of Conceptual Style.
- Low tolerance towards uncertainty and instinctive thinking way, are the characteristics of Behavioral Style.

From a realistic point of view, the most of the managers have dominant and alternative styles, some of them rely on their dominant style and some of them become flexible according to situations.

Social Responsibility of a Business

A business enterprise is an important part of the society and it should do its operations and earn money in ways that satisfy the expectations of the society. Social responsibility of a business refers to the obligations to take those decisions and perform those actions which are desirable in terms of the objectives and values of society. It is the idea that businesses should balance profit-making activities with activities that benefit society; it involves developing businesses with a positive relationship with the society in which they operate.

What is a business? The business itself is an economic activity. Its main function is to earn profits. To earn profits means to understand the needs and demands of consumers whether it be regarding the quality of the product or its price.

While understanding the perspective of the consumer and meeting their needs and demand to earn a profit is the economic responsibility of a business. When a business earns a profit, it also means that the employees earn the profit in terms of incentives. The economic growth of a business is not restricted to itself but affects the society as a whole.

Legal Responsibility

Legal responsibilities are not only liable to the individuals in the society but also to the businesses in the society. As business is an entity itself, it must also follow laws and rules. Every business has a responsibility to operate within the boundaries set by the various commissions and agencies at every level of the government. These rules and regulations are set for maintaining balance and

the greater good of the society. A law-abiding enterprise is a socially responsible enterprise as well. The business is free to do business however it wants but only within the boundaries of regulations of various laws such as labour law, environmental law and criminal law. For example, it's a business's duty to pay taxes to the government and keep its account books clean as it helps the government to track the economic state of the company.

Ethical Responsibility

Ethical responsibilities include the behaviour of the firm that is expected by the society but not codified in law. The factors of ethical responsibility include that the business must be environmentally friendly. The business should always be aware of its activities and how do they affect the environment. It is the moral and ethical responsibility of every human and every business.

Discretionary Responsibility

Business is one the most important pillar of the society. And therefore it should support and improve the society whenever it can. If a business is making significant profits it is the business responsibility that it should be philanthropic towards the society by donating funds or its goods and services.

It's the philanthropic responsibility of the business to help different groups of the society. It should also work towards providing free education by opening educational institutes and training institutes or helping the people affected by natural calamities such as floods and earthquakes. It is the responsibility of the company management to safeguard the capital investment by avoiding speculative activity and undertaking only healthy business ventures which give good returns on investment.

MEANING, NATURE AND SCOPE OF ECONOMICS

If we take a photograph of our country during day time, we shall find many of us engaged in various types of activities, most of which are related to earning one's own living and to the manner of satisfying one's own wants. These human activities, which are generally called economic, are mainly related to production,

distribution, consumption and exchange of goods and services. The scientific study of various problems arising out of these economic activities is called economics. These economic activities would not have been normally necessary if men did not want more than what they can produce.

But human nature being what it is, men do need more than they produce and hence there exists an economic problem. Human wants are many: the means to satisfy these wants are limited. These limited means are capable of alternative uses and hence there is a need of satisfying maximum wants with minimum of resources. The study of the attempts to do this is the subject matter of economics.

Definition of Business Economics

The teaching of economics is, thus, an abstract theorization with little connection to business. But theoretical models of economics are to be applied in business areas. Once theoretical models of economics are applied in business, the gap between economics and business gets minimised.

The branch of managerial economics or business economics has established links between business and economics. Business economics is, thus, an applied economics. Economics is the study of human beings (e.g., consumers, firms) in producing and consuming goods and services in the midst of scarcity of resources. Managerial or business economics is an applied branch of organising and allocating a firm's scarce resources to achieve its desired goals.

Managerial economics or business economics is economics applied in decision-making. Business economics, thus, interweaves economic principles and business. Business managers apply economic laws and principles while presenting business problems and their ways of solutions. Thus, business economics can be defined as the application of economic analysis to business problems faced by an enterprise. It provides a link between economic theory and the decision sciences in the analysis of managerial decision-making. It relies heavily on traditional economics and decision sciences.

Identification of the problems and the solving of the problems are the two crucial elements of decision-making of a business firm. Business economists help business managers in making sound business decisions. Business success, in fact, greatly depends on appropriate business decisions. However, appropriate decision-making is not an easy job in this changing world.

On the basis of past knowledge and experience, business managers take business decisions and make future plans. But decision-makers are constrained by the 'uncertainty' of the real world where changes occur either in a hidden way or in an open way. In this changing but uncertain world, an accurate decision-making is impossible even if talents of top quality business economists are employed.

It is due to this uncertainty, prediction or estimation relating to the volume of sales of a product, cost of production, profit, etc., is more likely to be imperfect. In other words, against the backdrop of uncertainty and a changing world, business managers will have to anticipate changes so that the impact of unfavorable situations becomes insignificant. Thus, business decision-making is an art.

Cultivation of this art is made through economic principles. In this sense, managerial economics is an applied economics. It is concerned with the application of economic concepts and analytical tools to the process of decision-making of a business enterprise.

Thus, managerial economics or business economics is a 'special branch of economics that bridges the gap between abstract economic theory and managerial practice. Through a process of application of the principles, concepts and tools of economics to solve the managerial problems of a business enterprise, business economists have greatly minimised the problem of uncertainty arising in business. The application field of economic theory is popularly known as business economics or managerial economics. Can economic theory be applied in business practice efficiently? Fritz Machlup, in answering this question, gave an analogy between the behaviour of a motorist deciding whether or not to overtake on a two-lane highway and the behaviour of a profit-maximising firm. Overtaking decision of the motorist involves construction of a very complex set of equations.

While overtaking, the motorist must have a knowledge about the weight, power, speed of the vehicle being driven, the condition of the road, weather, information about the number of vehicles plying on the highway, and a set of assumptions about the behaviour and objectives of other drivers. Unfortunately, even the most expert and cautious drivers do not have all these information.

But the model-builder makes such unrealistic assumptions. Under the circumstance, the decision of overtaking in a two-lane highway seems to be next to impossible. In reality, all the drivers overtake dozens of times every day. Such overtaking decisions are deemed to be 'correct' only if accidents are not met with.

If we compare this behavioural overtaking decision of a motorist with that of the behaviour of a profit-maximising firm, we will reach the same conclusion. A profit-maximising firm assumes that it has perfect information about costs and revenue conditions. But that too is an unrealistic assumption.

Still they assume such and are always guided by profit-maximising motive. "Like the driver taking the decision to overtake, managers 'behave' as if they had the relevant information, in which case they will behave like the profit-maximising model and that model will be a good predictor of their behaviour. "

Various Economic Methods

There are any numbers of methods which can be adopted. A country as a whole can adopt the collectivistic method in which one dictator will take decisions on behalf of all individuals. Or, there may be a completely free society in which all decisions are made by individuals. Actually there may be a combination of two methods with different emphasis on each of them. In fact economics is the study of the social effects of this decision making process. It decides which commodities and services are to be produced. It also allocates the available factors of production of each line of production. It decides for whom these goods and services have been produced.

Economics as Social Science

At present economics deals with a social phenomenon. We owe so many things to society such as division of labour, co-

operation of people in various economic activities, existence of money and its management, capital and technological knowledge, etc. And economics is nothing if not a study of all these things – money, capital, division of labour, etc. Economics is a social science because it is a study of human efforts to satisfy human wants in the conditions of limited resources which have alternative uses. These efforts and human activities are carried on in groups or in societies.

Economics and Human Behaviour

It is difficult to draw reliable generalizations about human behaviour in its ordinary business of life. Human behaviour is guided by so many impulses and is an outcome of various actions and reactions of social forces in which individual vagaries or individual whims and caprices, social laws and customs, traditions of a family or a group, political prejudices and emotions etc. play a part so that it is difficult to make any generalization. Thus, the first group of definitions seems to be defective in various ways, although it is not necessary to reject them totally. We should try to retain what is useful in them.

Economics and Scarcity

There are many wants which human beings would like to satisfy. Human beings try to satisfy as many of them as possible. Each of these wants is satiable, but taken together, one can say that wants are unlimited. Second, we have a large amount of resources capable of satisfying these wants. These resources are in fact means which are capable of alternative uses – although these resources are limited since they cannot be increased to any amount we desire.

Third, as wants are many and unlimited, while means are limited and capable of alternative uses, there is a need to economize. The ends have to be given varying importance and hence reason has to be exercised in choosing our ends or wants and in choosing the means for attaining them. The study of these choices and the generalizations which arise from the process of taking decisions is the subject matter of the science of economics.

Economics and Welfare

In making the study of economics wholly scientific we separated it from the human considerations. As a matter of fact, no social science can be studied in vacuum; it must take human beings into consideration, which will surely lead to the study of human welfare, and therefore to economic welfare. Second, when we talk of choice in this definition, it becomes too wide, because we are required to make choice for so many problems which might be political, biological, sociological, etc.

Its Relation with Other Social Sciences

There is no doubt about the fact that economics is very closely related to ethics. The early economic thinkers described economics as the handmaid of ethics. And formerly in the university teachings, economics was a part of moral sciences. Ethics is concerned with what ought to be. It describes the moral code of behaviour and is much more related to customs and traditions. At a glance we may be surprised as to how on earth economics can ever be related to ethics. Economics describes, while ethics prescribes. Economics is amoral, while ethics is moral. Economics is a positive science as it is concerned with 'what is', while ethics is a normative science as it is concerned with 'what ought to be'. Ethics is concerned with ends as such while economics is concerned with arranging means to ends; it has nothing to do with ends as such. Ethics is concerned with the mode of behaviour and the conduct of man while economics takes for granted all these things. Ethics is concerned with – all immaterial, eternal and sublime things while economics, many time, is concerned with all baser things in life – all material things.

Economics and Politics

The second very important field of social sciences with which economics is closely related is politics. One can almost say with confidence that it is difficult to draw a line between the two. They overlap each other. Politics is concerned with the philosophy of state and the methods of organizing the state. In both these respects economics has to enter into a relationship with politics. The

economic organization and the principles on which the economic system of a country works depend very much on the fundamental philosophy of the state. The nature of the working of the economic system differs very much in different types of states. Secondly, economics very often comes into closer relation with politics especially with the methods of organizing a state. In matters of legislations, like company law and nationalization, or in matters of public finance, economics plays a very important role.

In fact, nowadays, the government has become a huge economic machine. Most of the functions, such as, education, medical care, support of old people, unemployment, etc. which were formerly looked after by the church or family or group organizations, are now taken up by the state. A very important sector of the total national income is produced either through the state or within its control. Maintenance of full employment, optimum allocation of economic resources, equitable distribution of national income—all these have been the aims of the government economic policy, be it a socialist state or a capitalist state. It will not be surprising if the old term political economy will again come into use. Political economy is perhaps the most appropriate work for present day economics when planning is our main concern.

Economics and History

It is also closely related to history which is nothing but a scientific study of the records of facts. Mere facts cannot be history, when these facts are used in a particular way in order to bring out general trends, they constitute history. It is helpful in several ways. Economic facts of the past constitute the raw materials test tube, acid, powders etc. of a chemical laboratory for the economic historians, by the help of which he can test his economic theories. History of economic institutions, can also suggest the direction of change or evolution. As it is difficult to conduct controlled experiments in economics, history has a special importance.

At one time, the historical method occupied a very important place in the methods of economics. Moreover, the knowledge of history is essential for an economist who would like to apply his economic theory to the practical field. For the study of economic

frictions and to find out their importance, the study of history is very helpful; economic policy cannot be evolved without a close understanding of economic history.

Economics and Psychology

Similarly, economics is closely connected with psychology, mathematics and statistics. As economics, to a certain extent, is a study of choice or preferences, psychology comes into closer contact with it. Quite recently, psychology is developing into a science. There is a science of mass psychology. As division of labour has been adopted to a greater extent, man has no time to bother about anything else. He has become machine-like and hence there are reliable generalizations of mass psychology. Similarly we have industrial psychology, social psychology, etc. which help us a great deal in determining the economic quantities such as saving, investment, market demand curves, reactions to price changes, etc.

Economics, Statistics and Mathematics

Statistics is also very useful to economics. In fact, it has become a basis for economic studies. It helps an economist to verify his conclusions; it draws his attention to underlying currents. It helps him a great deal in interpreting economic facts of the past. No economist can do without statistics nowadays. Mathematics is also being increasingly used in economics with a view to sharpen economic analysis. Thus economics is closely connected with many other sciences.

Scope of Economics

The scope of Economics was extended to a large extent by Marshall and his views were supported by Pigou, Richards, Pension and Cannon. According to Pigou, "Economics is a study of economic welfare; economic welfare being described as that part of welfare which can be brought directly or indirectly into relations with the measuring rod of money."

According to Pension, "Economics is the science of material welfare." According to Cannon, "Economics is a study of the causes of material welfare."

There were many a critics of Marshall and they considered the definitions of Economics given by Marshall and his supporters to be unscientific. They were of the firm opinion that all these definitions the welfare conception of Economics lacked universality and scientific exactness. Robbins opines, "Economics is the science that studies human behaviour as a relationship between ends and scarce means which have alternative uses." The views of Stigler and Cairncross are also in agreement with those of Robbins.

Stigler observes, "Economics is the study of principles governing the allocation of scarce means among competing ends when the objective of allocation is to maximise the attainment of the ends."

In the opinion of Cairncross, "Economics is the study of influence of scarcity on human conduct in circumstances where men have freedom of choice in allocating scarce and competing wants." There are a large number of critics of Robbins, prominent among them being Durbin, Fraser, Wooten and Beveridge.

Wooten opines, "It is very difficult to, divest their discussions completely of all normative significance." In the views of Fraser, "Economics is something more than a value theory or equilibrium analysis."

Criticising the Robbins definition of Economics he critics observe that his definition misses the human touch which is of highest importance for study of any social science. Economics has been defined by Dr. J.K. Mehta, an eminent economist of India, as, "Economics is the science of human activities considered as an endeavour to reach the stage of wantlessness." This definition has grown more or less philosophically. It is believed that ultimately all the human wants will get fulfilled.

Though economics have not been able to come to any unanimity but most of them consider the definition given by Marshall as the best. Following generalisations which can be made by considering various definitions of economics helps us in understanding the scope of economics.

- (i) Economics is concerned mainly with the human activities but all the activities of man are not economic because some

of these can definitely be considered as religious, political, social etc. Such activities have no direct relationship with Economics. However in present day societies all the activities of man including social, political and religious are considered as economic activities because it is now believed that economic factors are responsible for all the activities of man.

- (ii) Economics considers man as a social being and is not concerned with the life of those persons who live away from the society.
- (iii) Economics concerns itself with the study of normal human beings and gives no place in its study to abnormal persons like lunatics, recluse etc.
- (iv) Economics tries to study man in an objective and scientific manner but makes use of literary language for presentation its studies. Thus it has elements of both science and literature. The aims and objects of Economics are based on science 1 its theories and other conclusions are more related to literature or art.
- (v) In Economics we are interested in the study or production and utility of the wealth. In its scope we also include the importance of the earning of wealth and its useful consumption or preservation. Following example is given to explain these thing

Economics, even if considered as a social science will suffer w some defects i.e. it concerns itself only with one aspect of human life, it is one sided only. The main aim of economics is the study of mate] aspect of human life. It tries to make all attempts to bring happiness man and to make him free of his wants. It tries to teach man the lesson of material utility. Thus, we find that the scope of Economics is very wide. As its c end we have human activities and at its other end we have the scientific and objective as also artistic study of the material things or objects.

Subject-matter of Economics

We have already seen that the scope of economics is very wide and so we can select the subject-matter from this wide field. Its

study cover' types of activities of human beings and includes their practical problem too. Supporting this point of view Pigou says, "When we study Economics our impulse is not the philosopher's impulse, knowledge for the sake of knowledge but rather the psychologists knowledge for the heading that knowledge may help to bring." Truly speaking Economics is a hand made of Ethics and a servant of practice. Looking at the Indian scene we find Government of India launching various plans and levying taxes to find money for execution of these plans. These taxes are being opposed by some persons. All these activities form a part of the subject-matter of Economics because all these activities have a direct or an indirect bearing on the economic aspect of man's life.

Division of Subject-matter

The subject-matter in Economics can be divided in following two parts:

- (i) Activities having a direct bearing on a single individual or a single family.
- (ii) Subjects having direct relationship with wider human circle like society, country, world.

Importance of Economics

In our country we are facing the problem of unemployment which is becoming acute for educated youth. It is therefore desirable that we pay more attention to and lay more emphasis on teaching those subjects that can help to solve the practical difficulties of life. In Economics we aim at teaching modern economic principles by observation and understanding of current practices. It should be our endure to train pupils to apply sound economic theory in their daily life.

Present Economic Structure

A fast change is taking place in the economic structure of different countries and it is essential as also useful to have a knowledge of the economic conditions of different countries. Economics helps us to gain such a knowledge and enables us to make it proper use. By studying Economics one becomes more

practical. Practical knowledge is quite essential for any one who wishes to acquire a thorough knowledge of Economics. Knowledge of Economics comes handy for solving various complex economic problems of the country.

To illustrate the fact that all the aspects of Economics have some practical values and importance. We take the following example Considering the prevalent state of poverty in most of Indian villages, foreign economists have drawn certain conclusions and such conclusions have been repeated by others without caring to find out the truth. The reasons for the poverty of Indian farmer have been put as under.

Indian farmer is without work for about six months in a year. However, this conclusion is definitely erroneous. If we look at the working of Indian farmer from a close angle we find that he is busy almost the whole year.

He is busy, during summers, in making arrangement for Rabi crops and if he finds any time during the period he utilises the same to perform various social functions such as marriages etc. Once the rainy season sets in he becomes busy to preparing his fields for the Kharif crop. By the time he is free from Kharif crop, the season for Rabi crop comes and in this way the cycle is repeated every year. It clearly illustrates that the cause for poverty of Indian farmer as given above is not correct and this conclusion has been drawn by foreign economists who do not have an intimate knowledge of Indian villages and the working conditions that exist there. This conclusion shifts the burden of cause of poverty from Government to the people. To properly analyse the situation and to thoroughly understand the conditions of rural areas, it is essential to have a thorough knowledge of rural economics.

If we try to analyse the causes for the poverty of Indian farmer in right perspective we can easily draw the following conclusions

- (i) Fragmentation of land.
- (ii) Inadequate means of irrigation.
- (iii) Use of orthodox methods of cultivation.
- (iv) Non-availability of improved seeds and good fertilisers due to scarcity of resources.

Thus, we find that practical knowledge is quite essential and such a knowledge is quite useful in teaching of Economics. More emphasis be put on teaching Economics with the things of daily use. It is desirable for teacher of Economics to always keep in mind the changing circumstances and conditions of the world.

Social Structure and Economic Structure

There occurs a change in the social structure whenever, there is a change in economic structure or we can simply say that socio-economic conditions are interdependent. To understand it, you can find for yourself by going to the market and finding out the price of a *Juggry*. The existing price is around Rs. 40 at present for the piece that was available for about Rs. 5 in 1947. Thus, the price has increased ten fold. Moreover, the item is not available at every place but can be purchased only from certain specified places. These changes can be attributed to change in social conditions.

Law of Demand and Supply

This is a very general law that works in all fields of life. We find an increase in demand for those things which are in greater need. The need for certain things increase or decrease with change in social structure. To illustrate it take the case of crystal sugar which was not so commonly used in Indian home about 25 years back and so was not much in demand. Now with change in social structure its demand has grown enormously not only in India but in the world itself. With this increase in demand of crystal sugar India took to export sugar and many large scale sugar factories were established. It affected the small scale industries which dwindled. At present the sugar prices are much more than what they used to be about 30 years back. This increase in sugar prices can be attributed to the law of demand and supply. In Economics one gets a knowledge of these things which he can use for the betterment of society and of the country.

Government Statistics

The statistics is issued, by the Government, periodically to highlight certain facts and achievements of the Government.

Generally this statistics is quite deceptive and to make a proper assessment of Government policies and programmes a good knowledge in economics is a must.

Budgeting

Budgeting occupies a very important place in our personal life as also in the affairs of a country. For an individual his personal budget is more important than the national budget. To lead a happy and contented life it is of prime importance to have a balanced family budget. This should be prepared in such a way that it is possible for us to meet our needs within our limited resources. All out efforts to made to save something for emergencies. For such a budgeting a knowledge of Economics is needed.

Practical Use in Mental Happiness

Mental happiness is definitely related to material happiness. Material happiness always brings with it the mental happiness. Economics can be put to practical use in life by drawing various charts, figures etc. These charts, diagrams etc., if drawn beautifully, they provide mental satisfaction. If these charts, diagrams etc., can provide us the knowledge of Economics then also we get pleasure and happiness.

Economics in the Knowledge of other Social Sciences

It is beyond doubt that the knowledge of Economics has a practical utility for the knowledge of Civics, History, Geography and other social sciences. A study of Economics provides us the knowledge of various resources and their practical utility. Knowledge of Economics has been used in making various new inventions and discoveries. Study of Economics helps us to know about various gains and losses. It is possible for us to draw various conclusions and theories on the basis of statistics available with us. Any country that has a good number of falls and rivers then these can be utilised to generate electricity which can be used to run various industries. This is a practical use of the knowledge of Economics.

Utility of Charts, Maps etc.

In teaching of Economics we frequently use charts and maps and they occupy an important place in the teaching process, while drawing these charts, maps an effort is also made to fill them with colours. This is done in an effort to provide a concrete form to certain abstract things. Such a presentation itself has a practical utility. For presenting these things in useful forms we need a knowledge of Economics.

From the above discussion it becomes quite clear that teaching of Economics has a practical value and importance.

CHARACTERISTICS OF BUSINESS ECONOMICS**Term Business Economics**

The term Business Economics is used in a variety of ways. Sometimes it is used as synonymously with-Industrial Economics-Industrial Organisation-Managerial Economics-Economics for Business. Industrial Economics is the mostly closely over-lapping of these terms whilst there may be more substantial differences with Economics for Business and Managerial Economics. One view of the distinctions between these would be that Business Economics is wider in its scope than Industrial Economics in that it would be concerned not only with “Industry” but also businesses in the service sector and that it also takes seriously the insights of the “business strategy” literature.

Economics for business looks at the major principles of economics but focuses on applying these economic principles to the real world of business. Managerial economics is the application of economic methods in the managerial decision-making process.

Characteristics of Business Economics

Business economics is essentially concerned with the various decisions of a business enterprise. The unit of study of business economics is the firm. Thus, managerial economics studies decision-making behaviour of a firm or an industry. Microeconomics takes into account the behaviour of smaller economic agents, such as a firm or a consumer or an input owner. It deals with the operation

of a consumer, a firm involving the determination of price of a commodity, revenue, costs and, hence, profit levels, etc. Managerial economics is, thus, essentially microeconomic in character as it has its origin in theoretical microeconomics. Profs. H. C. Petersen and W. C. Lewis suggest that managerial economics should be thought of as applied microeconomics.

It is an application of that part of microeconomics focusing on those topics which are of great interest and importance to business managers. These topics include theories of demand, production and cost, profit-maximising model of the firm, optimal prices and advertising expenditures, government regulation, etc. Managerial economics is concerned with finding optimal solutions to business decision problems.

Secondly, economic concepts and principles of the 'theory of firm' are employed in business economics. Thus, in business economics, the main emphasis is given upon the firm, the environment in which the firm finds itself, and the business decision which firms have to take. In this sense, managerial economics is narrower in scope than pure economic theory.

Thirdly, broadly there are two main branches of economics – 'positive' economics and 'normative' economics. Positive economics deals with 'description' while normative economics deals with 'prescription'. By building up propositions on the basis of a set of assumptions, positive economics tries to explain economic phenomenon.

Normative economics comments on the desirability of that phenomenon and suggests policy measures. Value judgments are, thus, pronounced in normative economics. In the words of Profs. Mote, Paul and Gupta: "Managerial economics is a part of normative economics as its focus is more on prescribing choice and action and less on explaining what has happened. Managerial economics draws on positive economics by utilizing the relevant theories as a basis for prescribing choices."

Fourthly, business economics not only seeks to investigate and analyse how and why businesses behave as they do but also the implications of their actions and policies for the industry in

which they operate and, finally, for the economy as a whole. In this business environment, both internal and external factors work. Business economics seeks to analyse various internal and external constraints that businesses experience in their process of growth and survival, draw conclusions as to how and why businesses behave as they do. "Business economics therefore focuses on the issues relevant to a business and its operations, and to the business environment." Thus, business economics is considered as applied economics. It casts away abstract economic theories. Managerial economists look at practical applications of theoretical models.

Finally, business economics is essentially microeconomic in character. In other words, macroeconomic theory has less relevance for managerial economics. Truly speaking, business economics should also deal with a wider environment – the macro-economy. Macro-economics is concerned with the behaviour of the economic system in totality. It studies the determination of aggregate national income, level of employment, general price level, the international balance of payments, etc.

It is true that aggregating economic trends or external economic factors do not directly affect business decisions of a firm. But what is true is that changing macro-economy not only influences aggregate or national income but also the demand for the product of a business firm. Efficient business managers must have awareness as well as keenness of studying and explaining macro- economic environment. In this sense, business economics cannot be devoid of macroeconomics.

Term Business Economics

The term Business Economics is used in a variety of ways. Sometimes it is used as synonymously with-Industrial Economics-Industrial Organisation-Managerial Economics-Economics for Business. Industrial Economics is the mostly closely over-lapping of these terms whilst there may be more substantial differences with Economics for Business and Managerial Economics. One view of the distinctions between these would be that Business Economics is wider in its scope than Industrial Economics in that it would be concerned not only with "Industry" but also businesses

in the service sector and that it also takes seriously the insights of the “business strategy” literature. Economics for business looks at the major principles of economics but focuses on applying these economic principles to the real world of business. Managerial economics is the application of economic methods in the managerial decision-making process.

Interpretations of Business Economics

Many universities offer courses in Business Economics and offer a range of interpretations as to the meaning of the term. The University of East London defines the subject matter of its degree as looking at the application of economic theory to business activities and organizations arguing that “In general terms, Business Economics deals with issues such as: the ways markets work; what firms do, what their motives are, how they perform; and the role of government in regulating business activity”. The program at Harvard University uses economic methods to analyse practical aspects of business, including business administration, management, and related fields of economics.

The University of Miami defines Business Economics as involving the study of how we use our resources for the production, distribution, and consumption of goods and services. This requires business economists to analyse social institutions, banks, the stock market, the government and they look at problems connected with labour negotiations, taxes, international trade, and urban and environmental issues. Courses at the University of Manchester interpret Business Economics to be concerned with the economic analysis of how businesses contribute to welfare of society rather than on the welfare of an individual or a business. This is done via an examination of the relationship between ownership, control and firm objectives; theories of the growth of the firm; the behavioural theory of the firm; theories of entrepreneurship; the factors that influence the structure, conduct and performance of business at the industry level.

Business Process Outsourcing

Business process outsourcing (BPO) is a subset of outsourcing that involves the contracting of the operations and responsibilities

of specific business functions (or processes) to a third-party service provider. Originally, this was associated with manufacturing firms, such as Coca Cola that outsourced large segments of its supply chain. In the contemporary context, it is primarily used to refer to the outsourcing of business processing services to an outside firm, replacing in-house services with labour from an outside firm.

BPO is typically categorized into back office outsourcing-which includes *internal business functions* such as human resources or finance and accounting, and front office outsourcing-which includes *customer-related services* such as contact Centre services. BPO that is contracted outside a company's country is called offshore outsourcing. BPO that is contracted to a company's neighboring (or nearby) country is called nearshore outsourcing.

Given the proximity of BPO to the information technology industry, it is also categorized as an information technology enabled service or ITES. Knowledge process outsourcing (KPO) and legal process outsourcing (LPO) are some of the sub-segments of business process outsourcing industry.

Subject-Matter and Scope of Business Economics

We know that managerial economics or business economics is applied microeconomics employed for the purpose of facilitating decision--making and forward planning. As far as decision--making and forward planning are concerned, one has to face the following problems:

Problem of resource allocation seems to be a pressing problem for any organisation. Resources are not plentiful. A firm has to organise scarce resources efficiently so that optimal outcomes are obtained. Such resource allocation problem includes production programming, transportation problem, etc. Non-optimal organisation of resources may spell disaster to any organisation.

Inventory and queuing are important problems to any firm. A firm has to hold an optimal level of stocks of raw materials and finished product so that business uncertainties can be minimised. Business managers must decide an optimal level of inventories. Such decisions are taken by firms after considering demand and

supply conditions.

Since forward planning by management is essential, a firm must make decisions – whether new machines are to be installed or more professionals are to be employed. As most of the decisions cannot be implemented simultaneously, the firm manager must make a trade-off between decisions. Taking a particular decision out of a variety of decisions is known as queuing problem. A manager places or queues alternative decisions and picks up a right one. Price fixation is another interrelated problem connected with decision-making. A firm has to take up a right pricing decision.

Finally, the decision-maker faces investment problems for a variety of reasons. Truly speaking, any forward planning by management involves investment problems which are by nature knotty. Investment problems boil down to the problems of allocating resources over time. A firm has to make decision about the volume of investment. It must decide where to invest, when to invest. It must know the sources of funds, etc.

Anyway, business economics is concerned with decision-making and forward planning. The scope of business economics (both micro and macro variety) is a wider one since it “uses the logic of Economics, Mathematics and Statistics to provide effective ways of thinking about business decision problems.” In view of this saying of Prof. D. C. Hague, we can argue that there are links between managerial economics and management science. In fact, the boundaries between the two subjects are not clear-cut but overlapping.

Managerial economics is largely an applied branch of microeconomics. Its macroeconomic content is not to be belittled. It uses the methods and techniques of microeconomics mostly in the field of management. As Haynes and William Warren state: “The relation of managerial economics to economic theory (of either the micro or macro variety) is much like that of engineering to physics, or of medicine to biology or bacteriology. It is the relation of an applied field to the more fundamental but more abstract basic discipline from which it borrows concepts and analytical tools.”

It is to be pointed out here that measurement without theory may lead to false precision and diagnosis while theory without measurement can hardly be operationally useful. Now we are in a position to explain the scope of business economics. By scope of (business) economics we mean the field of the subject, the boundaries that delimit and delineate the topics to be addressed.

Determination of the scope of the subject includes:

- (i) Definition of the subject,
- (ii) Subject matter of business economics,
- (iii) Is business economics a positive or a normative science?

Since business economics is thought of as applied microeconomics, the scope of business economics includes:

- i. Analysis, estimation and forecasting of consumer demand for a product;
- ii. Analysis of cost and output;
- iii. Determination of price of a commodity, work policy and business strategies of a business enterprise;
- iv. Long-run planning decisions of a firm that studies capital budgeting and capital management.

Managerial economics now includes 'operations research' – a mathematical technique to solve business problems.

Finally, we must say that there is a great deal of linkages between managerial economics and other disciplines and fields of study. It uses the logic of economics, mathematics and statistics. Managerial economics is related to management science or the decision sciences. Management science is concerned with techniques for improvement of decision-making.

Business economics is related to accounting. Accounting is essentially concerned with recording and analysing the financial activities of a business firm. To quote Prof. D. C. Hague: "The main task of management accounting is now seen as being to provide the sort of data which managers need if they are to apply the idea of managerial economics to solve business problems correctly; accounting data are also to be provided in a form so as to fit easily into the concepts and analysis of managerial economics."

In view of the relationship between managerial economics and other disciplines, it may be called an art, and not a science. Though an art, decision-making in this uncertain world has become more perfect. Better choices, better prediction, etc., are likely to emerge because of the interaction among basic disciplines.

Thus, it is clear from the above discussion that managerial or business economics helps managers of firms, administrators of non-profit and profit-making hospitals, schools, colleges and universities to recognise how economic forces affect organisations. It applies economic theory and methods to business and administrative decision-making in both profit and non-profit sector.

Scope of Business Economics

Business Economics covers most of the problems that a manager or establishment faces. Hence, the scope of business economics is wide. Since a firm can face internal/operational as well as external and/or environmental issues, there are different economic theories applicable to them. Microeconomics helps with internal or operational issues whereas macroeconomics is applied to external or environmental issues. In this article, we will look at the scope of business economics under both these heads.

Microeconomics Applied to Operational Issues

As the name suggests, internal or operational issues are issues that arise within a firm and are within the control of the management. It is within the scope of business economics to analyze this.

Further, a few examples of such issues are choice of business, size of business, product designs, pricing, promotion for sales, technology choice, etc. Most firms can deal with these using the following microeconomics theories:

Analyzing Demand and Forecasting

Analyzing demand is all about understanding buyer behavior. It studies the preferences of consumers along with the effects of changes in the determinants of demand. Also, these determinants include the price of the good, consumer's income, tastes/

preferences, etc. Forecasting demand is a technique used to predict the future demand for a good and/or service. Further, this prediction is based on the past behavior of factors which affect the demand. This is important for firms as accurate predictions help them produce the required quantities of goods at the right time.

Further, it gives them enough time to arrange various factors of production in advance like raw materials, labor, equipment, etc. Business Economics offers scientific tools which assist in forecasting demand.

NATIONAL INCOME AND ITS IMPORTANCE IN BUSINESS ECONOMICS

Concept of National Income

The National Income is the total amount of income accruing to a country from economic activities in a years time. It includes payments made to all resources either in the form of wages, interest, rent, and profits.

The progress of a country can be determined by the growth of the national income of the country

National Income Definition

There are two National Income Definition

- Traditional Definition
- Modern Definition

Traditional Definition

According to Marshall: "The labor and capital of a country acting on its natural resources produce annually a certain net aggregate of commodities, material and immaterial including services of all kinds. This is the true net annual income or revenue of the country or national dividend."

The definition as laid down by Marshall is being criticized on the following grounds. Due to the varied category of goods and services, a correct estimation is very difficult. There is a chance of double counting, hence National Income cannot be estimated correctly.

For example, a product runs in the supply from the producer to distributor to wholesaler to retailer and then to the ultimate consumer. If on every movement commodity is taken into consideration then the value of National Income increases.

Also, one other reason is that there are products which are produced but not marketed. For example, In an agriculture-oriented country like India, there are commodities which though produced but are kept for self-consumption or exchanged with other commodities. Thus there can be an underestimation of National Income. Simon Kuznets defines national income as “the net output of commodities and services flowing during the year from the country’s productive system in the hands of the ultimate consumers.”

Following are the Modern National Income definition

- GDP
- GNP

Gross Domestic Product

The total value of goods produced and services rendered within a country during a year is its Gross Domestic Product.

Further, GDP is calculated at market price and is defined as GDP at market prices. Different constituents of GDP are:

1. Wages and salaries
2. Rent
3. Interest
4. Undistributed profits
5. Mixed-income
6. Direct taxes
7. Dividend
8. Depreciation

Gross National Product

For calculation of GNP, we need to collect and assess the data from all productive activities, such as agricultural produce, wood, minerals, commodities, the contributions to production by

transport, communications, insurance companies, professions such (as lawyers, doctors, teachers, etc). at market prices.

It also includes net income arising in a country from abroad. Four main constituents of GNP are:

1. Consumer goods and services
2. Gross private domestic income
3. Goods produced or services rendered
4. Income arising from abroad.

Methods of Measuring National Income

There are four methods of measuring national income. Which method is to be used depends on the availability of data in a country and the purpose in hand.

- (1) Product Method: According to this method, the total value of final goods and services produced in a country during a year is calculated at market prices. To find out the GNP, the data of all productive activities, such as agricultural products, wood received from forests, minerals received from mines, commodities produced by industries, the contributions to production made by transport, communications, insurance companies, lawyers, doctors, teachers, etc. are collected and assessed at market prices. Only the final goods and services are included and the intermediary goods and services are left out.
- (2) Income Method: According to this method, the net income payments received by all citizens of a country in a particular year are added up, i.e., net incomes that accrue to all factors of production by way of net rents, net wages, net interest and net profits are all added together but incomes received in the form of transfer payments are not included in it. The data pertaining to income are obtained from different sources, for instance, from income tax department in respect of high income groups and in case of workers from their wage bills.
- (3) Expenditure Method: According to this method, the total expenditure incurred by the society in a particular year is

added together and includes personal consumption expenditure, net domestic investment, government expenditure on goods and services, and net foreign investment. This concept is based on the assumption that national income equals national expenditure.

- (4) Value Added Method: Another method of measuring national income is the value added by industries. The difference between the value of material outputs and inputs at each stage of production is the value added. If all such differences are added up for all industries in the economy, we arrive at the gross domestic product.

Difficulties or Limitations in Measuring National Income

There are many conceptual and statistical problems involved in measuring national income by the income method, product method, and expenditure method.

We discuss them separately in the light of the three methods:

Problems in Income Method

The following problems arise in the computation of National Income by income method:

1. Owner-occupied Houses: A person who rents a house to another earns rental income, but if he occupies the house himself, will the services of the house-owner be included in national income. The services of the owner-occupied house are included in national income as if the owner sells to himself as a tenant its services.

For the purpose of national income accounts, the amount of imputed rent is estimated as the sum for which the owner-occupied house could have been rented. The imputed net rent is calculated as that portion of the amount that would have accrued to the house-owner after deducting all expenses.

2. Self-employed Persons: Another problem arises with regard to the income of self-employed persons. In their case, it is very difficult to find out the different inputs provided by

the owner himself. He might be contributing his capital, land, labour and his abilities in the business. But it is not possible to estimate the value of each factor input to production. So he gets a mixed income consisting of interest, rent, wage and profits for his factor services. This is included in national income.

3. Goods meant for Self-consumption: In under-developed countries like India, farmers keep a large portion of food and other goods produced on the farm for self-consumption. The problem is whether that part of the produce which is not sold in the market can be included in national income or not. If the farmer were to sell his entire produce in the market, he will have to buy what he needs for self-consumption out of his money income. If, instead he keeps some produce for his self-consumption, it has money value which must be included in national income.
4. Wages and Salaries paid in Kind: Another problem arises with regard to wages and salaries paid in kind to the employees in the form of free food, lodging, dress and other amenities. Payments in kind by employers are included in national income. This is because the employees would have received money income equal to the value of free food, lodging, etc. from the employer and spent the same in paying for food, lodging, etc.

Problems in Product Method

The following problems arise in the computation of national income by product method:

Services of Housewives

The estimation of the unpaid services of the housewife in the national income presents a serious difficulty. A housewife renders a number of useful services like preparation of meals, serving, tailoring, mending, washing, cleaning, bringing up children, etc.

She is not paid for them and her services are not including in national income. Such services performed by paid servants are

included in national income. The national income is, therefore, underestimated by excluding the services of a housewife. The reason for the exclusion of her services from national income is that the love and affection of a housewife in performing her domestic work cannot be measured in monetary terms. That is why when the owner of a firm marries his lady secretary, her services are not included in national income when she stops working as a secretary and becomes a housewife. When a teacher teaches his own children, his work is also not included in national income. Similarly, there are a number of goods and services which are difficult to be assessed in money terms for the reason stated above, such as painting, singing, dancing, etc. as hobbies.

Intermediate and Final Goods

The greatest difficulty in estimating national income by product method is the failure to distinguish properly between intermediate and final goods. There is always the possibility of including a good or service more than once, whereas only final goods are included in national income estimates. This leads to the problem of double counting which leads to the overestimation of national income.

Second-hand Goods and Assets

Another problem arises with regard to the sale and purchase of second-hand goods and assets. We find that old scooters, cars, houses, machinery, etc. are transacted daily in the country. But they are not included in national income because they were counted in the national product in the year they were manufactured.

If they are included every time they are bought and sold, national income would increase many times. Similarly, the sale and purchase of old stocks, shares, and bonds of companies are not included in national income because they were included in national income when the companies were started for the first time. Now they are simply financial transactions and represent claims. But the commission or fees charged by the brokers in the repurchase and resale of old shares, bonds, houses, cars or scooters, etc. are included in national income. For these are the payments they receive for their productive services during the year.

Illegal Activities

Income earned through illegal activities like gambling, smuggling, illicit extraction of wine, etc. is not included in national income. Such activities have value and satisfy the wants of the people but they are not considered productive from the point of view of society. But in countries like Nepal and Monaco where gambling is legalised, it is included in national income. Similarly, horse-racing is a legal activity in England and is included in national income.

Consumers' Service

There are a number of persons in society who render services to consumers but they do not produce anything tangible. They are the actors, dancers, doctors, singers, teachers, musicians, lawyers, barbers, etc. The problem arises about the inclusion of their services in national income since they do not produce tangible commodities. But as they satisfy human wants and receive payments for their services, their services are included as final goods in estimating national income.

Capital Gains

The problem also arises with regard to capital gains. Capital gains arise when a capital asset such as a house, some other property, stocks or shares, etc. is sold at higher price than was paid for it at the time of purchase. Capital gains are excluded from national income because these do not arise from current economic activities. Similarly, capital losses are not taken into account while estimating national income.

Inventory Changes

All inventory changes (or changes in stocks) whether positive or negative are included in national income. The procedure is to take changes in physical units of inventories for the year valued at average current prices paid for them. The value of changes in inventories may be positive or negative which is added or subtracted from the current production of the firm. Remember, it is the change in inventories and not total inventories for the year that are taken into account in national income estimates.

Depreciation

Depreciation is deducted from GNP in order to arrive at NNP. Thus depreciation lowers the national income. But the problem is of estimating the current depreciated value of, say, a machine, whose expected life is supposed to be thirty years. Firms calculate the depreciation value on the original cost of machines for their expected life. This does not solve the problem because the prices of machines change almost every year.

Price Changes

National income by product method is measured by the value of final goods and services at current market prices. But prices do not remain stable. They rise or fall. When the price level rises, the national income also rises, though the national production might have fallen.

On the contrary, with the fall in the price level, the national income also falls, though the national production might have increased. So price changes do not adequately measure national income. To solve this problem, economists calculate the real national income at a constant price level by the consumer price index.

Problems in Expenditure Method

The following problems arise in the calculation of national income by expenditure method:

Government Services

In calculating national income by, expenditure method, the problem of estimating government services arises. Government provides a number of services, such as police and military services, administrative and legal services. Should expenditure on government services be included in national income?

If they are final goods, then only they would be included in national income. On the other hand, if they are used as intermediate goods, meant for further production, they would not be included in national income. There are many divergent views on this issue.

One view is that if police, military, legal and administrative services protect the lives, property and liberty of the people, they

are treated as final goods and hence form part of national income. If they help in the smooth functioning of the production process by maintaining peace and security, then they are like intermediate goods that do not enter into national income.

In reality, it is not possible to make a clear demarcation as to which service protects the people and which protects the productive process. Therefore, all such services are regarded as final goods and are included in national income.

Transfer Payments

There arises the problem of including transfer payments in national income. Government makes payments in the form of pensions, unemployment allowance, subsidies, interest on national debt, etc. These are government expenditures but they are not included in national income because they are paid without adding anything to the production process during the current year.

For instance, pensions and unemployment allowances are paid to individuals by the government without doing any productive work during the year. Subsidies tend to lower the market price of the commodities. Interest on national or public debt is also considered a transfer payment because it is paid by the government to individuals and firms on their past savings without any productive work.

Durable-use Consumers' Goods

Durable-use consumers' goods also pose a problem. Such durable-use consumers' goods as scooters, cars, fans, TVs, furniture's, etc. are bought in one year but they are used for a number of years. Should they be included under investment expenditure or consumption expenditure in national income estimates? The expenditure on them is regarded as final consumption expenditure because it is not possible to measure their used up value for the subsequent years.

But there is one exception. The expenditure on a new house is regarded as investment expenditure and not consumption expenditure. This is because the rental income or the imputed rent which the house-owner gets is for making investment on the new

house. However, expenditure on a car by a household is consumption expenditure. But if he spends the amount for using it as a taxi, it is investment expenditure.

Public Expenditure

Government spends on police, military, administrative and legal services, parks, street lighting, irrigation, museums, education, public health, roads, canals, buildings, etc. The problem is to find out which expenditure is consumption expenditure and which investment expenditure is. Expenses on education, museums, public health, police, parks, street lighting, civil and judicial administration are consumption expenditure. Expenses on roads, canals, buildings, etc. are investment expenditure. But expenses on defence equipment are treated as consumption expenditure because they are consumed during a war as they are destroyed or become obsolete. However, all such expenses including the salaries of armed personnel are included in national income.

Importance of National Income Analysis

The national income data have the following importance:

For the Economy

National income data are of great importance for the economy of a country. These days the national income data are regarded as accounts of the economy, which are known as social accounts. These refer to net national income and net national expenditure, which ultimately equal each other.

Social accounts tell us how the aggregates of a nation's income, output and product result from the income of different individuals, products of industries and transactions of international trade. Their main constituents are inter-related and each particular account can be used to verify the correctness of any other account.

National Policies

National income data form the basis of national policies such as employment policy, because these figures enable us to know the direction in which the industrial output, investment and

savings, etc. change, and proper measures can be adopted to bring the economy to the right path.

Economic Planning

In the present age of planning, the national data are of great importance. For economic planning, it is essential that the data pertaining to a country's gross income, output, saving and consumption from different sources should be available. Without these, planning is not possible.

Economic Models

The economists propound short-run as well as long-run economic models or long-run investment models in which the national income data are very widely used.

Research

The national income data are also made use of by the research scholars of economics. They make use of the various data of the country's input, output, income, saving, consumption, investment, employment, etc., which are obtained from social accounts.

Per Capita Income

National income data are significant for a country's per capita income which reflects the economic welfare of the country. The higher the per capita income, the higher the economic welfare of the country.

Distribution of Income

National income statistics enable us to know about the distribution of income in the country. From the data pertaining to wages, rent, interest and profits, we learn of the disparities in the incomes of different sections of the society. Similarly, the regional distribution of income is revealed. It is only on the basis of these that the government can adopt measures to remove the inequalities in income distribution and to restore regional equilibrium. With a view to removing these personal and regional disequilibria, the decisions to levy more taxes and increase public expenditure also rest on national income statistics.

Alternative Methods Used for Measuring National Income

The three alternative methods used for measuring national income are as follows: 1. Value Added Method 2. Income Method 3. Expenditure Method.

Since factor incomes arise from the production of goods and services, and since incomes are spent on goods and services produced, three alternative methods of measuring national income are possible.

Value Added Method

This is also called output method or production method. In this method the value added by each enterprise in the production goods and services is measured. Value added by an enterprise is obtained by deducting expenditure incurred on intermediate goods such as raw materials, unfinished goods (purchased from other firms from the value of output produced by an enterprise.

Value of output produced by an enterprise is equal to physical output (Q) produced multiplied by the market price (P), that is, $P \cdot Q$. From the value added by each enterprise we subtract consumption of fixed capital (i.e., depreciation) to obtain net value added at market prices (NVA_{MP}).

However, for estimating national income (that is, Net National Product at factor cost (NNP_{FC})) we require to estimate net value added at factor cost (NVA_{FC}) by each enterprise in the economy. NVA_{FC} can be found out by deducting net indirect taxes (i. e. indirect taxes less subsidies provided by the Government).

Under this method, the economy is divided into different industrial sectors such as agriculture, fishing, mining, construction, manufacturing, trade and commerce, transport, communication and other services. Then, the net value added at factor cost (NVA_{FC}) by each productive enterprise as well as by each industry or sector is estimated.

It follows from above that in order to arrive at the net value added at factor cost by an enterprise we have to subtract the following from the value of output of an enterprise:

1. Intermediate consumption which is the value of goods such as raw materials, fuels purchased from other firms
2. Consumption of fixed capital (i.e., depreciation)
3. Net indirect taxes.

Summing up the net values added at factor cost (NVA_{FC}) by all productive enterprises of an industry or sector gives us the net value added at factor cost of each industry or sector. We then add up net values added at factor cost by all industries or sectors to get net domestic product at factor cost (NDP_{FC}). Lastly, to the net domestic product we add the net factor income from abroad to get net national product at factor cost (NNP_{FC}) which is also called national income. Thus,

$$NI \text{ or } NNP_{FC} = NDP_{FC} + \text{Net factor income from abroad}$$

This method of calculating national income can be used where there exists a census of production for the year. In many countries, the data of production of only important industries are known. Hence this method is employed along with other methods to arrive at the national income. The one great advantage of this method is that it reveals the relative importance of the different sectors of the economy by showing their respective contributions to the national income.

Precautions

The following precautions should be taken while measuring national income of a country through value added method:

1. Imputed rent values of self-occupied houses should be included in the value of output. Though these payments are not made to others, their values can be easily estimated from prevailing values in the market.
2. Sale and purchase of second-hand goods should not be included in measuring value of output of a year because their values were counted in the year of output of the year of their production. Of course, commission or brokerage earned in their sale and purchase has to be included because this is a new service rendered in the current year.
3. Value of production for self-consumption are be counted

while measuring national income. In this method, the production for self-consumption should be valued at the prevailing market prices.

4. Value of services of housewives are not included because it is not easy to find out correctly the value of their services.
5. Value of intermediate goods must not be counted while measuring value added because this will amount to double counting.

Income Method

This method approaches national income from distribution side. In other words, this method measures national income at the phase of distribution and appears as income paid and or received by individuals of the country. Thus, under this method, national income is obtained by summing up of the incomes of all individuals of a country. Individuals earn incomes by contributing their own services and the services of their property such as land and capital to the national production.

Therefore, national income is calculated by adding up the rent of land, wages and salaries of employees, interest on capital, profits of entrepreneurs (including undistributed corporate profits) and incomes of self-employed people. This method of estimating national income has the great advantage of indicating the distribution of national income among different income groups such as landlords, owners of capital, workers, entrepreneurs.

Measurement of national income through income method involves the following main steps:

1. Like the value added method, the first step in income method is also to identify the productive enterprises and then classify them into various industrial sectors such as agriculture, fishing, forestry, manufacturing, transport, trade and commerce, banking, etc.
2. The second step is to classify the factor payments. The factor payments are classified into the following groups:
 - i. Compensation of employees which includes wages and salaries, both in cash and kind, as well as employers' contribution to social security schemes.

- ii. Rent and also royalty, if any.
- iii. Interest.
- iv. Profits:

Profits are divided into three sub-groups:

- (i) Dividends
- (ii) Undistributed profits
- (iii) Corporate income tax
- v. Mixed income of the self-employed:

In India as in other developing countries there is fifth category of factor income which is termed as mixed income of self-employed. In India a good number of people are engaged in household industries, in family farms and other unorganised enterprises. Because of self-employment nature of the business it is difficult to separate wages for the work done by the self-employed from the surplus or profits made by them. Therefore, the incomes earned by them are mix of wages, rent, interest and profit and are, therefore, called mixed income of the self-employed.

3. The third step is to measure factor payments. Income paid out by each enterprise can be estimated by gathering information about the number of units of each factor employed and the income paid out to each unit of every factor. Price paid out to each factor multiplied by the number of units of each factor employed would give us the factor's income.
4. The adding up of factor payments by all enterprises belonging to an industrial sector would give us the incomes paid out to various factors by a particular industrial sector.
5. By summing up the incomes paid out by all industrial sectors we will obtain domestic factor income which is also called net domestic product at factor cost (NDP_{FC}).
6. Finally, by adding net factor income earned from abroad to domestic factor income or NDP_{FC} we get net national product at factor cost (NNP_{FC}) which is also called national income.

Precautions

While estimating national income through income method the following precautions should be taken:

1. Transfer payments are not included in estimating national income through this method.
2. Imputed rent of self-occupied houses are included in national income as these houses provide services to those who occupy them and its value can be easily estimated from the market value data.
3. Illegal money such as hawala money, money earned through smuggling etc. are not included as they cannot be easily estimated.
4. Windfall gains such as prizes won, lotteries are also not included.
5. Corporate profit tax (that is, tax on income of the companies) should not be separately included as it has already been included as a part of profits.
6. Death duties, gift tax, wealth tax, tax on lotteries, etc., are paid from past savings or wealth and not from current income. Therefore, they should not be treated as a part of national income of a year.
7. The receipts from the sale of second-hand goods should not be treated as a part of national income. This is because the sale of second-hand goods does not create new flows goods and services in the current year.
8. Income equal to the value of production used for self-consumption should be estimated and included in the measure of national income.

Expenditure Method

Expenditure method arrives at national income by adding up all expenditures made on goods and services during a year. Income can be spent either on consumer goods or capital goods. Again, expenditure can be made by private individuals and households or by government and business enterprises. Further, people of foreign countries spend on the goods and services which a country

exports to them. Similarly, people of a country spend on imports of goods and services from other countries. We add up the following types of expenditure by households, government and by productive enterprises to obtain national income.

1. Expenditure on consumer goods and services by individuals and households. This is called final private consumption expenditure, and is denoted by C.
2. Government's expenditure on goods and services to satisfy collective wants. This is called government's final consumption expenditure, and is denoted by G.
3. The expenditure by productive enterprises on capital goods and inventories or stocks. This is called gross domestic-capital formation, or gross domestic investment and is denoted by I or GDCF.

Gross domestic capital formation is divided into two parts:

- (i) Gross fixed capital formation
 - (ii) Addition to the stocks or inventories of goods
4. The expenditure made by foreigners on goods and services of a country exported to other countries which are called exports and are denoted by X. We deduct from exports (X) the expenditure by people, enterprises and government of a country on imports (M) of goods and services from other countries. That is, we have to estimate net exports (that is, exports - imports) or $(X - M)$ which is also denoted by NX.

Thus, we add up the above four types of expenditure to get final expenditure on gross domestic product at market prices (GDP_{MP}). Thus,

GDP_{MP} = Private final consumption expenditure + Government's final consumption expenditure + Gross domestic capital formation + Exports – Imports or

$$\begin{aligned} GDP_{MP} &= C + G + I + (X - M) \\ &= C + G + I + NX \end{aligned}$$

On deducting consumption of fixed capital (i.e., depreciation) from gross domestic product at market prices (GDP_{MP}) we get net domestic product at market prices (NDP_{MP}).

Lastly, we add 'net factor income from abroad' to obtain net national product at factor cost (NNP_{FC}), which is called national income. Thus,

$NNP_{FC} = GDP_{MP} - \text{Consumption of Fixed capital} - \text{Net Indirect taxes} + \text{Net Factor Income from Abroad.}$

Precautions

While estimating Gross Domestic Product through expenditure method or measuring final expenditure on Gross National Product, the following precautions should be taken:

1. Second-hand goods: The expenditure made on second-hand goods should not be included because this does not contribute to the current year production of goods and services.
2. Purchase of shares and bonds: Expenditure on purchase of old shares and bonds from other people and from business enterprises should not be included while estimating Gross Domestic Product through expenditure method. This is because bonds and shares are mere financial claims and do not represent expenditure on currently produced goods and services.
3. Expenditure on transfer payments by government such as unemployment benefits, old-age pension should also not be included because no goods or productive services are produced in exchange by the recipients of these payments.
4. Expenditure on intermediate goods such as fertilisers and seeds by the farmers and wool, cotton and yarn by manufacturers of garments should also be excluded. This is because we have to avoid double counting. Therefore, for estimating Gross Domestic Product we have to include only expenditure on final goods and services.

A greatest difficulty in the measurement of national income in the developing countries is general lack of adequate statistical data. Inadequacy, non-availability and unreliability of statistics is a great handicap in measuring national income in these countries.

Statistical information regarding agriculture and allied occupations, and household enterprises is not available. Even the

statistical information regarding the enterprises in the organised sector is sketchy and unreliable. There is no accurate information available regarding consumption, investment expenditure and savings of either rural or urban population.

GDP and GNP on the basis of Market Price and Factor Cost

Market Price

The Actual transacted price including indirect taxes such as GST, Customs duty etc. Such taxes tend to raise the prices of goods and services in the economy.

Factor Cost

It Includes the cost of factors of production e.g. interest on capital, wages to labor, rent for land profit to the stakeholders. Thus services provided by service providers and goods sold by the producer is equal to revenue price.

Alternatively,

Revenue Price (or Factor Cost) = Market Price (net of) Net Indirect Taxes

Net Indirect Taxes = Indirect Taxes Net of Subsidies received

Hence,

Factor Cost shall be equal to

(Market Price) LESS (Indirect Taxes ADD Subsidies)

Net Domestic Product

The net output of the country's economy during a year is its NDP. During the year a country's capital assets are subject to wear and tear due to its use or can become obsolete.

Hence, we deduct a percentage of such investment from the GDP to arrive at NDP.

So $NDP = GDP \text{ at factor cost LESS Depreciation.}$

The Accumulation of all factors of income earned by residents of a country and includes income earned from the county as well as from abroad.

Thus, National Income at Factor Cost shall be equal to
NNP at Market Price LESS (Indirect Taxes ADD Subsidies)

Measures of national income and output

A variety of measures of national income and output are used in economics to estimate total economic activity in a country or region, including gross domestic product (GDP), gross national product (GNP), net national income (NNI), and adjusted national income (NNI adjusted for natural resource depletion – also called as NNI at factor cost). All are specially concerned with counting the total amount of goods and services produced within the economy and by different sectors. The boundary is usually defined by geography or citizenship, and it is also defined as the total income of the nation and also restrict the goods and services that are counted. For instance, some measures count only goods & services that are exchanged for money, excluding bartered goods, while other measures may attempt to include bartered goods by *imputing* monetary values to them.

National accounts

The total production of goods and services in a large region like a country entails a large amount of data-collection and calculation. Although some attempts were made to estimate national incomes as long ago as the 17th century, the systematic keeping of national accounts, of which these figures are a part, only began in the 1930s, in the United States and some European countries.

The impetus for that major statistical effort was the Great Depression and the rise of Keynesian economics, which prescribed a greater role for the government in managing an economy, and made it necessary for governments to obtain accurate information so that their interventions into the economy could proceed as well-informed as possible.

Market value

In order to count a good or service, it is necessary to assign value to it. The value that the measures of national income and output assign to a good or service is its market value – the price

it fetches when bought or sold. The actual usefulness of a product (its use-value) is not measured – assuming the use-value to be any different from its market value.

Three strategies have been used to obtain the market values of all the goods and services produced: the product (or output) method, the expenditure method, and the income method. The product method looks at the economy on an industry-by-industry basis.

The total output of the economy is the sum of the outputs of every industry. However, since an output of one industry may be used by another industry and become part of the output of that second industry, to avoid counting the item twice we use not the value output by each industry, but the value-added; that is, the difference between the value of what it puts out and what it takes in. The total value produced by the economy is the sum of the values-added by every industry.

The expenditure method is based on the idea that all products are bought by somebody or some organisation. Therefore, we sum up the total amount of money people and organisations spend in buying things.

This amount must equal the value of everything produced. Usually, expenditures by private individuals, expenditures by businesses, and expenditures by government are calculated separately and then summed to give the total expenditure. Also, a correction term must be introduced to account for imports and exports outside the boundary.

The income method works by summing the incomes of all producers within the boundary. Since what they are paid is just the market value of their product, their total income must be the total value of the product. Wages, proprietor's incomes, and corporate profits are the major subdivisions of income.

Methods of measuring national income

Output

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

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

Key formulae are:

GDP(gross domestic product) at market price = value of output in an economy in the particular year minus intermediate consumption

GDP at factor cost = GDP at market price minus depreciation plus NFIA (*net factor income from abroad*) minus net indirect taxes(GNP)

NDP at factor cost = Compensation of employees plus net interest plus rental & royalty income plus profit of incorporated and unincorporated NDP at factor cost

The names of the measures consist of one of the words "Gross" or "Net", followed by one of the words "National" or "Domestic", followed by one of the words "Product", "Income", or "Expenditure". All of these terms can be explained separately.

"Gross" means total product, regardless of the use to which it is subsequently put.

"Net" means "Gross" minus the amount that must be used to offset depreciation – ie., wear-and-tear or obsolescence of the nation's fixed capital assets. "Net" gives an indication of how much product is actually available for consumption or new investment.

"Domestic" means the boundary is geographical: we are counting all goods and services produced within the country's

borders, regardless of by whom. "National" means the boundary is defined by citizenship (nationality). We count all goods and services produced by the nationals of the country (or businesses owned by them) regardless of where that production physically takes place.

The output of a French-owned cotton factory in Senegal counts as part of the Domestic figures for Senegal, but the National figures of France.

"Product", "Income", and "Expenditure" refer to the three counting methodologies explained earlier: the product, income, and expenditure approaches. However the terms are used loosely.

"Product" is the general term, often used when any of the three approaches was actually used. Sometimes the word "Product" is used and then some additional symbol or phrase to indicate the methodology; so, for instance, we get "Gross Domestic Product by income", "GDP (income)", "GDP(I)", and similar constructions.

"Income" specifically means that the income approach was used.

"Expenditure" specifically means that the expenditure approach was used.

Note that all three counting methods should in theory give the same final figure. However, in practice minor differences are obtained from the three methods for several reasons, including changes in inventory levels and errors in the statistics. One problem for instance is that goods in inventory have been produced (therefore included in Product), but not yet sold (therefore not yet included in Expenditure). Similar timing issues can also cause a slight discrepancy between the value of goods produced (Product) and the payments to the factors that produced the goods (Income), particularly if inputs are purchased on credit, and also because wages are collected often after a period of production.

National Income: Definition, Concepts and Methods of Measuring National Income

National income is an uncertain term which is used interchangeably with national dividend, national output and

national expenditure. On this basis, national income has been defined in a number of ways. In common parlance, national income means the total value of goods and services produced annually in a country. In other words, the total amount of income accruing to a country from economic activities in a year's time is known as national income. It includes payments made to all resources in the form of wages, interest, rent and profits.

Definitions of National Income

The definitions of national income can be grouped into two classes: One, the traditional definitions advanced by Marshall, Pigou and Fisher; and two, modern definitions.

The Marshallian Definition

According to Marshall: "The labour and capital of a country acting on its natural resources produce annually a certain net aggregate of commodities, material and immaterial including services of all kinds. This is the true net annual income or revenue of the country or national dividend." In this definition, the word 'net' refers to deductions from the gross national income in respect of depreciation and wearing out of machines. And to this, must be added income from abroad.

Its Defects: Though the definition advanced by Marshall is simple and comprehensive, yet it suffers from a number of limitations. First, in the present day world, so varied and numerous are the goods and services produced that it is very difficult to have a correct estimation of them.

Consequently, the national income cannot be calculated correctly. Second, there always exists the fear of the mistake of double counting, and hence the national income cannot be correctly estimated. Double counting means that a particular commodity or service like raw material or labour, etc. might get included in the national income twice or more than twice.

For example, a peasant sells wheat worth Rs.2000 to a flour mill which sells wheat flour to the wholesaler and the wholesaler sells it to the retailer who, in turn, sells it to the customers. If each time, this wheat or its flour is taken into consideration, it will work

out to Rs.8000, whereas, in actuality, there is only an increase of Rs.2000 in the national income.

Third, it is again not possible to have a correct estimation of national income because many of the commodities produced are not marketed and the producer either keeps the produce for self-consumption or exchanges it for other commodities. It generally happens in an agriculture- oriented country like India. Thus the volume of national income is underestimated.

The Pigouvian Definition

A.C. Pigou has in his definition of national income included that income which can be measured in terms of money. In the words of Pigou, "National income is that part of objective income of the community, including of course income derived from abroad which can be measured in money."

This definition is better than the Marshallian definition. It has proved to be more practical also. While calculating the national income now-a- days, estimates are prepared in accordance with the two criteria laid down in this definition.

First, avoiding double counting, the goods and services which can be measured in money are included in national income. Second, income received on account of investment in foreign countries is included in national income.

It's Defects: The Pigouvian definition is precise, simple and practical but it is not free from criticism. First, in the light of the definition put forth by Pigou, we have to unnecessarily differentiate between commodities which can and which cannot be exchanged for money.

But, in actuality, there is no difference in the fundamental forms of such commodities, no matter they can be exchanged for money. Second, according to this definition when only such commodities as can be exchanged for money are included in estimation of national income, the national income cannot be correctly measured.

According to Pigou, a woman's services as a nurse would be included in national income but excluded when she worked in the

home to look after her children because she did not receive any salary for it. Similarly, Pigou is of the view that if a man marries his lady secretary, the national income diminishes as he has no longer to pay for her services.

Thus the Pigovian definition gives rise to a number of paradoxes. Third, the Pigovian definition is applicable only to the developed countries where goods and services are exchanged for money in the market.

According to this definition, in the backward and underdeveloped countries of the world, where a major portion of the produce is simply bartered, correct estimate of national income will not be possible, because it will always work out less than the real level of income. Thus the definition advanced by Pigou has a limited scope.

Fisher's Definition

Fisher adopted 'consumption' as the criterion of national income whereas Marshall and Pigou regarded it to be production. According to Fisher, "The National dividend or income consists solely of services as received by ultimate consumers, whether from their material or from the human environments. Thus, a piano, or an overcoat made for me this year is not a part of this year's income, but an addition to the capital. Only the services rendered to me during this year by these things are income."

Fisher's definition is considered to be better than that of Marshall or Pigou, because Fisher's definition provides an adequate concept of economic welfare which is dependent on consumption and consumption represents our standard of living.

It's Defects: But from the practical point of view, this definition is less useful, because there are certain difficulties in measuring the goods and services in terms of money. First, it is more difficult to estimate the money value of net consumption than that of net production.

In one country there are several individuals who consume a particular good and that too at different places and, therefore, it is very difficult to estimate their total consumption in terms of

money. Second, certain consumption goods are durable and last for many years.

If we consider the example of piano or overcoat, as given by Fisher, only the services rendered for use during one year by them will be included in income.

If an overcoat costs Rs. 100 and lasts for ten years, Fisher will take into account only Rs. 100 as national income during one year, whereas Marshall and Pigou will include Rs. 100 in the national income for the year, when it is made.

Besides, it cannot be said with certainty that the overcoat will last only for ten years. It may last longer or for a shorter period. Third, the durable goods generally keep changing hands leading to a change in their ownership and value too.

It, therefore, becomes difficult to measure in money the service-value of these goods from the point of view of consumption. For instance, the owner of a Maruti car sells it at a price higher than its real price and the purchaser after using it for a number of years further sells it at its actual price.

Now the question is as to which of its price, whether actual or black market one, should we take into account, and afterwards when it is transferred from one person to another, which of its value according to its average age should be included in national income?

But the definitions advanced by Marshall, Pigou and Fisher are not altogether flawless. However, the Marshallian and Pigovian definitions tell us of the reasons influencing economic welfare, whereas Fisher's definition helps us compare economic welfare in different years.

Modern Definitions

From the modern point of view, Simon Kuznets has defined national income as "the net output of commodities and services flowing during the year from the country's productive system in the hands of the ultimate consumers."

On the other hand, in one of the reports of United Nations, national income has been defined on the basis of the systems of

estimating national income, as net national product, as addition to the shares of different factors, and as net national expenditure in a country in a year's time. In practice, while estimating national income, any of these three definitions may be adopted, because the same national income would be derived, if different items were correctly included in the estimate.

Concepts of National Income

There are a number of concepts pertaining to national income and methods of measurement relating to them.

Gross Domestic Product (GDP)

GDP is the total value of goods and services produced within the country during a year. This is calculated at market prices and is known as GDP at market prices. Dernberg defines GDP at market price as "the market value of the output of final goods and services produced in the domestic territory of a country during an accounting year."

There are three different ways to measure GDP:

Product Method, Income Method and Expenditure Method.

These three methods of calculating GDP yield the same result because National Product = National Income = National Expenditure.

1. The Product Method: In this method, the value of all goods and services produced in different industries during the year is added up. This is also known as the value added method to GDP or GDP at factor cost by industry of origin. The following items are included in India in this: agriculture and allied services; mining; manufacturing, construction, electricity, gas and water supply; transport, communication and trade; banking and insurance, real estates and ownership of dwellings and business services; and public administration and defense and other services (or government services). In other words, it is the sum of gross value added.
2. The Income Method: The people of a country who produce GDP during a year receive incomes from their work. Thus

GDP by income method is the sum of all factor incomes: Wages and Salaries (compensation of employees) + Rent + Interest + Profit.

3. Expenditure Method: This method focuses on goods and services produced within the country during one year.

GDP by expenditure method includes:

- (1) Consumer expenditure on services and durable and non-durable goods (C),
- (2) Investment in fixed capital such as residential and non-residential building, machinery, and inventories (I),
- (3) Government expenditure on final goods and services (G),
- (4) Export of goods and services produced by the people of country (X),
- (5) Less imports (M). That part of consumption, investment and government expenditure which is spent on imports is subtracted from GDP. Similarly, any imported component, such as raw materials, which is used in the manufacture of export goods, is also excluded.

Thus GDP by expenditure method at market prices = $C + I + G + (X - M)$, where $(X - M)$ is net export which can be positive or negative.

GDP at Factor Cost

GDP at factor cost is the sum of net value added by all producers within the country. Since the net value added gets distributed as income to the owners of factors of production, GDP is the sum of domestic factor incomes and fixed capital consumption (or depreciation).

Thus GDP at Factor Cost = Net value added + Depreciation.

GDP at factor cost includes:

- (i) Compensation of employees i.e., wages, salaries, etc.
- (ii) Operating surplus which is the business profit of both incorporated and unincorporated firms. [Operating Surplus = Gross Value Added at Factor Cost – Compensation of Employees – Depreciation]

(iii) Mixed Income of Self- employed.

Conceptually, GDP at factor cost and GDP at market price must be identical/This is because the factor cost (payments to factors) of producing goods must equal the final value of goods and services at market prices. However, the market value of goods and services is different from the earnings of the factors of production.

In GDP at market price are included indirect taxes and are excluded subsidies by the government. Therefore, in order to arrive at GDP at factor cost, indirect taxes are subtracted and subsidies are added to GDP at market price.

Thus, $\text{GDP at Factor Cost} = \text{GDP at Market Price} - \text{Indirect Taxes} + \text{Subsidies}$.

Net Domestic Product (NDP)

NDP is the value of net output of the economy during the year. Some of the country's capital equipment wears out or becomes obsolete each year during the production process. The value of this capital consumption is some percentage of gross investment which is deducted from GDP. Thus $\text{Net Domestic Product} = \text{GDP at Factor Cost} - \text{Depreciation}$.

Nominal and Real GDP

When GDP is measured on the basis of current price, it is called GDP at current prices or nominal GDP. On the other hand, when GDP is calculated on the basis of fixed prices in some year, it is called GDP at constant prices or real GDP.

Nominal GDP is the value of goods and services produced in a year and measured in terms of rupees (money) at current (market) prices. In comparing one year with another, we are faced with the problem that the rupee is not a stable measure of purchasing power. GDP may rise a great deal in a year, not because the economy has been growing rapidly but because of rise in prices (or inflation).

On the contrary, GDP may increase as a result of fall in prices in a year but actually it may be less as compared to the last year. In both 5 cases, GDP does not show the real state of the economy.

To rectify the underestimation and overestimation of GDP, we need a measure that adjusts for rising and falling prices.

This can be done by measuring GDP at constant prices which is called real GDP. To find out the real GDP, a base year is chosen when the general price level is normal, i.e., it is neither too high nor too low. The prices are set to 100 (or 1) in the base year.

Now the general price level of the year for which real GDP is to be calculated is related to the base year on the basis of the following formula which is called the deflator index:

$$\text{Real GDP} = \frac{\text{GDP for the Current Year}}{\frac{\text{Base Year (=100)}}{\text{Current Year Index}}}$$

Suppose 1990-91 is the base year and GDP for 1999-2000 is Rs. 6, 00,000 crores and the price index for this year is 300.

Thus, Real GDP for 1999-2000 = Rs. 6, 00,000 × 100/300 = Rs. 2, 00,000 crores

GDP Deflator

GDP deflator is an index of price changes of goods and services included in GDP. It is a price index which is calculated by dividing the nominal GDP in a given year by the real GDP for the same year and multiplying it by 100. Thus,

$$\text{GDP Deflator} = \frac{\text{Nominal (or Current Prices) GDP}}{\text{Real (or Constant Prices) GDP}} \times 100$$

For example, GDP Deflator in 1997-98 = $\frac{1426.7 \text{ th. crores}}{1049.2 \text{ th. crores at } 1993-94} \times 100 = 135.9$

It shows that at constant prices (1993-94), GDP in 1997-98 increased by 135.9% due to inflation (or rise in prices) from Rs. 1049.2 thousand crores in 1993-94 to Rs. 1426.7 thousand crores in 1997-98.

Gross National Product (GNP)

GNP is the total measure of the flow of goods and services at market value resulting from current production during a year in a country, including net income from abroad.

GNP includes four types of final goods and services:

- (1) Consumers' goods and services to satisfy the immediate wants of the people;
- (2) Gross private domestic investment in capital goods consisting of fixed capital formation, residential construction and inventories of finished and unfinished goods;
- (3) Goods and services produced by the government; and
- (4) Net exports of goods and services, i.e., the difference between value of exports and imports of goods and services, known as net income from abroad.

In this concept of GNP, there are certain factors that have to be taken into consideration: First, GNP is the measure of money, in which all kinds of goods and services produced in a country during one year are measured in terms of money at current prices and then added together.

But in this manner, due to an increase or decrease in the prices, the GNP shows a rise or decline, which may not be real. To guard against erring on this account, a particular year (say for instance 1990-91) when prices be normal, is taken as the base year and the GNP is adjusted in accordance with the index number for that year. This will be known as GNP at 1990-91 prices or at constant prices.

Second, in estimating GNP of the economy, the market price of only the final products should be taken into account. Many of the products pass through a number of stages before they are ultimately purchased by consumers.

If those products were counted at every stage, they would be included many a time in the national product. Consequently, the GNP would increase too much. To avoid double counting, therefore, only the final products and not the intermediary goods should be taken into account.

Third, goods and services rendered free of charge are not included in the GNP, because it is not possible to have a correct estimate of their market price. For example, the bringing up of a child by the mother, imparting instructions to his son by a teacher, recitals to his friends by a musician, etc.

Fourth, the transactions which do not arise from the produce of current year or which do not contribute in any way to production are not included in the GNP. The sale and purchase of old goods, and of shares, bonds and assets of existing companies are not included in GNP because these do not make any addition to the national product, and the goods are simply transferred.

Fifth, the payments received under social security, e.g., unemployment insurance allowance, old age pension, and interest on public loans are also not included in GNP, because the recipients do not provide any service in lieu of them. But the depreciation of machines, plants and other capital goods is not deducted from GNP.

Sixth, the profits earned or losses incurred on account of changes in capital assets as a result of fluctuations in market prices are not included in the GNP if they are not responsible for current production or economic activity.

For example, if the price of a house or a piece of land increases due to inflation, the profit earned by selling it will not be a part of GNP. But if, during the current year, a portion of a house is constructed anew, the increase in the value of the house (after subtracting the cost of the newly constructed portion) will be included in the GNP. Similarly, variations in the value of assets, that can be ascertained beforehand and are insured against flood or fire, are not included in the GNP.

Last, the income earned through illegal activities is not included in the GNP. Although the goods sold in the black market are priced and fulfill the needs of the people, but as they are not useful from the social point of view, the income received from their sale and purchase is always excluded from the GNP.

There are two main reasons for this. One, it is not known whether these things were produced during the current year or the preceding years. Two, many of these goods are foreign made and smuggled and hence not included in the GNP.

Three Approaches to GNP

After having studied the fundamental constituents of GNP, it is essential to know how it is estimated. Three approaches are

employed for this purpose. One, the income method to GNP; two, the expenditure method to GNP and three, the value added method to GNP. Since gross income equals gross expenditure, GNP estimated by all these methods would be the same with appropriate adjustments.

Income Method to GNP

The income method to GNP consists of the remuneration paid in terms of money to the factors of production annually in a country. Thus GNP is the sum total of the following items:

- (i) Wages and salaries: Under this head are included all forms of wages and salaries earned through productive activities by workers and entrepreneurs. It includes all sums received or deposited during a year by way of all types of contributions like overtime, commission, provident fund, insurance, etc.
- (ii) Rents: Total rent includes the rents of land, shop, house, factory, etc. and the estimated rents of all such assets as are used by the owners themselves.
- (iii) Interest: Under interest comes the income by way of interest received by the individual of a country from different sources. To this is added, the estimated interest on that private capital which is invested and not borrowed by the businessman in his personal business. But the interest received on governmental loans has to be excluded, because it is a mere transfer of national income.
- (iv) Dividends: Dividends earned by the shareholders from companies are included in the GNP.
- (v) Undistributed corporate profits: Profits which are not distributed by companies and are retained by them are included in the GNP.
- (vi) Mixed incomes: These include profits of unincorporated business, self-employed persons and partnerships. They form part of GNP.
- (vii) Direct taxes: Taxes levied on individuals, corporations and other businesses are included in the GNP.

- (viii) Indirect taxes: The government levies a number of indirect taxes, like excise duties and sales tax.
These taxes are included in the price of commodities. But revenue from these goes to the government treasury and not to the factors of production. Therefore, the income due to such taxes is added to the GNP.
- (ix) Depreciation: Every corporation makes allowance for expenditure on wearing out and depreciation of machines, plants and other capital equipment. Since this sum also is not a part of the income received by the factors of production, it is, therefore, also included in the GNP.
- (x) Net income earned from abroad: This is the difference between the value of exports of goods and services and the value of imports of goods and services. If this difference is positive, it is added to the GNP and if it is negative, it is deducted from the GNP.

Thus GNP according to the Income Method = Wages and Salaries + Rents + Interest + Dividends + Undistributed Corporate Profits + Mixed Income + Direct Taxes + Indirect Taxes + Depreciation + Net Income from abroad.

Expenditure Method to GNP

From the expenditure view point, GNP is the sum total of expenditure incurred on goods and services during one year in a country.

It includes the following items:

- (i) Private consumption expenditure: It includes all types of expenditure on personal consumption by the individuals of a country. It comprises expenses on durable goods like watch, bicycle, radio, etc., expenditure on single-used consumers' goods like milk, bread, ghee, clothes, etc., as also the expenditure incurred on services of all kinds like fees for school, doctor, lawyer and transport. All these are taken as final goods.
- (ii) Gross domestic private investment: Under this comes the expenditure incurred by private enterprise on new investment and on replacement of old capital. It includes

expenditure on house construction, factory- buildings, and all types of machinery, plants and capital equipment.

In particular, the increase or decrease in inventory is added to or subtracted from it. The inventory includes produced but unsold manufactured and semi-manufactured goods during the year and the stocks of raw materials, which have to be accounted for in GNP. It does not take into account the financial exchange of shares and stocks because their sale and purchase is not real investment. But depreciation is added.

- (iii) Net foreign investment: It means the difference between exports and imports or export surplus. Every country exports to or imports from certain foreign countries. The imported goods are not produced within the country and hence cannot be included in national income, but the exported goods are manufactured within the country. Therefore, the difference of value between exports (X) and imports (M), whether positive or negative, is included in the GNP.
- (iv) Government expenditure on goods and services: The expenditure incurred by the government on goods and services is a part of the GNP. Central, state or local governments spend a lot on their employees, police and army. To run the offices, the governments have also to spend on contingencies which include paper, pen, pencil and various types of stationery, cloth, furniture, cars, etc.

It also includes the expenditure on government enterprises. But expenditure on transfer payments is not added, because these payments are not made in exchange for goods and services produced during the current year.

Thus GNP according to the Expenditure Method = Private Consumption Expenditure (C) + Gross Domestic Private Investment (I) + Net Foreign Investment (X-M) + Government Expenditure on Goods and Services (G) = $C + I + (X-M) + G$.

As already pointed out above, GNP estimated by either the income or the expenditure method would work out to be the same, if all the items are correctly calculated.

Value Added Method to GNP

Another method of measuring GNP is by value added. In calculating GNP, the money value of final goods and services produced at current prices during a year is taken into account. This is one of the ways to avoid double counting. But it is difficult to distinguish properly between a final product and an intermediate product.

For instance, raw materials, semi-finished products, fuels and services, etc. are sold as inputs by one industry to the other. They may be final goods for one industry and intermediate for others. So, to avoid duplication, the value of intermediate products used in manufacturing final products must be subtracted from the value of total output of each industry in the economy.

Thus, the difference between the value of material outputs and inputs at each stage of production is called the value added. If all such differences are added up for all industries in the economy, we arrive at the GNP by value added. $\text{GNP by value added} = \text{Gross value added} + \text{net income from abroad}$.

The supposition that the entire economy for purposes of total production consists of three sectors. They are agriculture, manufacturing, and others, consisting of the tertiary sector.

Out of the value of total output of each sector is deducted the value of its intermediate purchases (or primary inputs) to arrive at the value added for the entire economy.

Thus the value of total output of the entire economy is Rs. 155 crores and the value of its primary inputs comes to Rs. 80 crores. Thus the GDP by value added is Rs. 75 crores (Rs. 155 minus Rs. 80 crores).

The total value added equals the value of gross domestic product of the economy. Out of this value added, the major portion goes in the form wages and salaries, rent, interest and profits, a small portion goes to the government as indirect taxes and the remaining amount is meant for depreciation.

Thus we find that the total gross value added of an economy equals the value of its gross domestic product. If depreciation is

deducted from the gross value added, we have net value added which comes to Rs. 67 crores (Rs. 75 minus Rs. 8 crores).

This is nothing but net domestic product at market prices. Again, if indirect taxes (Rs. 7 crores) are deducted from the net domestic product of Rs. 67 crores, we get Rs. 60 crores as the net value added at factor cost which is equivalent to net domestic product at factor cost.

It's Importance: The value added method for measuring national income is more realistic than the product and income methods because it avoids the problem of double counting by excluding the value of intermediate products.

Thus this method establishes the importance of intermediate products in the national economy.

Second, by studying the national income accounts relating to value added, the contribution of each production sector to the value of the GNP can be found out.

For instance, it can tell us whether agriculture is contributing more or the share of manufacturing is falling, or of the tertiary sector is increasing in the current year as compared to some previous years. Third, this method is highly useful because "it provides a means of checking the GNP estimates obtained by summing the various types of commodity purchases."

It's Difficulties: However, difficulties arise in the calculation of value added in the case of certain public services like police, military, health, education, etc. which cannot be estimated accurately in money terms. Similarly, it is difficult to estimate the contribution made to value added by profits earned on irrigation and power projects.

GNP at Market Prices

When we multiply the total output produced in one year by their market prices prevalent during that year in a country, we get the Gross National Product at market prices. Thus GNP at market prices means the gross value of final goods and services produced annually in a country plus net income from abroad. It includes the gross value of output of all items from (1) to (4)

mentioned under GNP. $\text{GNP at Market Prices} = \text{GDP at Market Prices} + \text{Net Income from Abroad}$.

GNP at Factor Cost

GNP at factor cost is the sum of the money value of the income produced by and accruing to the various factors of production in one year in a country. It includes all items mentioned above under income method to GNP less indirect taxes.

GNP at market prices always includes indirect taxes levied by the government on goods which raise their prices. But GNP at factor cost is the income which the factors of production receive in return for their services alone. It is the cost of production.

Thus GNP at market prices is always higher than GNP at factor cost. Therefore, in order to arrive at GNP at factor cost, we deduct indirect taxes from GNP at market prices. Again, it often happens that the cost of production of a commodity to the producer is higher than a price of a similar commodity in the market.

In order to protect such producers, the government helps them by granting monetary help in the form of a subsidy equal to the difference between the market price and the cost of production of the commodity. As a result, the price of the commodity to the producer is reduced and equals the market price of similar commodity. For example if the market price of rice is Rs. 3 per kg but it costs the producers in certain areas Rs. 3.50. The government gives a subsidy of 50 paise per kg to them in order to meet their cost of production. Thus in order to arrive at GNP at factor cost, subsidies are added to GNP at market prices.

$\text{GNP at Factor Cost} = \text{GNP at Market Prices} - \text{Indirect Taxes} + \text{Subsidies}$.

Net National Product (NNP)

NNP includes the value of total output of consumption goods and investment goods. But the process of production uses up a certain amount of fixed capital. Some fixed equipment wears out, its other components are damaged or destroyed, and still others are rendered obsolete through technological changes.

All this process is termed depreciation or capital consumption allowance. In order to arrive at NNP, we deduct depreciation from GNP. The word 'net' refers to the exclusion of that part of total output which represents depreciation. So $NNP = GNP - \text{Depreciation}$.

NNP at Market Prices

Net National Product at market prices is the net value of final goods and services evaluated at market prices in the course of one year in a country. If we deduct depreciation from GNP at market prices, we get NNP at market prices. So $NNP \text{ at Market Prices} = GNP \text{ at Market Prices} - \text{Depreciation}$.

NNP at Factor Cost

Net National Product at factor cost is the net output evaluated at factor prices. It includes income earned by factors of production through participation in the production process such as wages and salaries, rents, profits, etc. It is also called National Income. This measure differs from NNP at market prices in that indirect taxes are deducted and subsidies are added to NNP at market prices in order to arrive at NNP at factor cost. Thus

$NNP \text{ at Factor Cost} = NNP \text{ at Market Prices} - \text{Indirect taxes} + \text{Subsidies}$

$= GNP \text{ at Market Prices} - \text{Depreciation} - \text{Indirect taxes} + \text{Subsidies}$.

$= \text{National Income}$.

Normally, NNP at market prices is higher than NNP at factor cost because indirect taxes exceed government subsidies. However, NNP at market prices can be less than NNP at factor cost when government subsidies exceed indirect taxes.

Domestic Income

Income generated (or earned) by factors of production within the country from its own resources is called domestic income or domestic product.

Domestic income includes: (i) Wages and salaries, (ii) rents,

including imputed house rents, (iii) interest, (iv) dividends, (v) undistributed corporate profits, including surpluses of public undertakings, (vi) mixed incomes consisting of profits of unincorporated firms, self-employed persons, partnerships, etc., and (vii) direct taxes.

Since domestic income does not include income earned from abroad, it can also be shown as: Domestic Income = National Income - Net income earned from abroad.

Thus the difference between domestic income and national income is the net income earned from abroad. If we add net income from abroad to domestic income, we get national income, i.e., National Income = Domestic Income + Net income earned from abroad.

But the net national income earned from abroad may be positive or negative. If exports exceed imports, net income earned from abroad is positive. In this case, national income is greater than domestic income. On the other hand, when imports exceed exports, net income earned from abroad is negative and domestic income is greater than national income.

Private Income

Private income is income obtained by private individuals from any source, productive or otherwise, and the retained income of corporations. It can be arrived at from NNP at Factor Cost by making certain additions and deductions.

The additions include transfer payments such as pensions, unemployment allowances, sickness and other social security benefits, gifts and remittances from abroad, windfall gains from lotteries or from horse racing, and interest on public debt. The deductions include income from government departments as well as surpluses from public undertakings, and employees' contribution to social security schemes like provident funds, life insurance, etc.

Thus Private Income = National Income (or NNP at Factor Cost) + Transfer Payments + Interest on Public Debt – Social Security – Profits and Surpluses of Public Undertakings.

Personal Income

Personal income is the total income received by the individuals of a country from all sources before payment of direct taxes in one year. Personal income is never equal to the national income, because the former includes the transfer payments whereas they are not included in national income. Personal income is derived from national income by deducting undistributed corporate profits, profit taxes, and employees' contributions to social security schemes. These three components are excluded from national income because they do reach individuals.

But business and government transfer payments, and transfer payments from abroad in the form of gifts and remittances, windfall gains, and interest on public debt which are a source of income for individuals are added to national income. Thus Personal Income = National Income - Undistributed Corporate Profits - Profit Taxes - Social Security Contribution + Transfer Payments + Interest on Public Debt. Personal income differs from private income in that it is less than the latter because it excludes undistributed corporate profits.

Thus Personal Income = Private Income - Undistributed Corporate Profits - Profit Taxes.

Disposable Income

Disposable income or personal disposable income means the actual income which can be spent on consumption by individuals and families. The whole of the personal income cannot be spent on consumption, because it is the income that accrues before direct taxes have actually been paid. Therefore, in order to obtain disposable income, direct taxes are deducted from personal income. Thus Disposable Income = Personal Income - Direct Taxes.

But the whole of disposable income is not spent on consumption and a part of it is saved. Therefore, disposable income is divided into consumption expenditure and savings. Thus Disposable Income = Consumption Expenditure + Savings.

If disposable income is to be deduced from national income, we deduct indirect taxes plus subsidies, direct taxes on personal and on business, social security payments, undistributed corporate

profits or business savings from it and add transfer payments and net income from abroad to it.

Thus Disposable Income = National Income – Business Savings – Indirect Taxes + Subsidies – Direct Taxes on Persons – Direct Taxes on Business – Social Security Payments + Transfer Payments + Net Income from abroad.

Real Income

Real income is national income expressed in terms of a general level of prices of a particular year taken as base. National income is the value of goods and services produced as expressed in terms of money at current prices. But it does not indicate the real state of the economy.

It is possible that the net national product of goods and services this year might have been less than that of the last year, but owing to an increase in prices, NNP might be higher this year. On the contrary, it is also possible that NNP might have increased but the price level might have fallen, as a result national income would appear to be less than that of the last year. In both the situations, the national income does not depict the real state of the country. To rectify such a mistake, the concept of real income has been evolved.

In order to find out the real income of a country, a particular year is taken as the base year when the general price level is neither too high nor too low and the price level for that year is assumed to be 100. Now the general level of prices of the given year for which the national income (real) is to be determined is assessed in accordance with the prices of the base year. For this purpose the following formula is employed.

Real NNP = $\text{NNP for the Current Year} \times \text{Base Year Index} / 100$ / Current Year Index

Suppose 1990-91 is the base year and the national income for 1999-2000 is Rs. 20,000 crores and the index number for this year is 250. Hence, Real National Income for 1999-2000 will be = $20000 \times 100 / 250$ = Rs. 8000 crores. This is also known as national income at constant prices.

Demand Analysis

DETERMINANTS OF DEMAND

Determinants

Availability of raw materials: for example, availability may cap the amount of gold that can be produced in a country regardless of price. Likewise, the price of Van Gogh paintings is unlikely to affect their supply. Length and complexity of production: Much depends on the complexity of the production process. Textile production is relatively simple. The labour is largely unskilled and production facilities are little more than buildings – no special structures are needed. Thus the PES for textiles is elastic. On the other hand, the PES for specific types of motor vehicles is relatively inelastic. Auto manufacture is a multi-stage process that requires specialized equipment, skilled labour, a large suppliers network and large R&D costs.

Time to respond: The more time a producer has to respond to price changes the more elastic the supply. Supply is normally more elastic in the long run than in the short run for produced goods, since it is generally assumed that in the long run all factors of production can be utilised to increase supply, whereas in the short run only labour can be increased, and even then, changes may be prohibitively costly. For example, a cotton farmer cannot immediately (i.e. in the short run) respond to an increase in the price of soybeans because of the time it would take to procure the necessary land. Excess capacity: A producer who has unused

capacity can (and will) quickly respond to price changes in his market assuming that variable factors are readily available.

Inventories: A producer who has a supply of goods or available storage capacity can quickly increase supply to market. Various research methods are used to calculate price elasticities in real life, including analysis of historic sales data, both public and private, and use of present-day surveys of customers' preferences to build up test markets capable of modelling such changes. Alternatively, conjoint analysis (a ranking of users' preferences which can then be statistically analysed) may be used.

Determinants of Demand with Examples and Formula

Demand drives economic growth. Businesses want to increase demand so they can improve profits. Governments and central banks boost demand to end recessions. They slow it during the expansion phase of the business cycle to combat inflation. If you offer any paid services, even you are trying to raise demand for them.

What drives demand? In economics, there are five determinants of individual demand and a sixth for aggregate demand.

The Five Determinants of Demand

The five determinants of demand are:

1. The price of the good or service.
2. Income of buyers.
3. Prices of related goods or services. These are either complementary, those purchased along with a particular good or service, or substitutes, those purchased instead of a certain good or service.
4. Tastes or preferences of consumers.
5. Expectations. These are usually about whether the price will go up.

For aggregate demand, the number of buyers in the market is the sixth determinant.

Demand Equation or Function

This equation expresses the relationship between demand and its five determinants:

$q_D = f$ (price, income, prices of related goods, tastes, expectations)

It says that the quantity demanded of a product is a function of five factors: price, income of the buyer, the price of related goods, the tastes of the consumer, and any expectation the consumer has of future supply, prices, etc.

How Each Determinant Affects Demand

You can understand how each determinant affects demand if you first assume that all the other determinants don't change. That principle is called *ceteris paribus* or "all other things being equal." So, *ceteris paribus*, here's how each element affects demand.

Price. The law of demand states that when prices rise, the quantity of demand falls. That also means that when prices drop, demand will grow. People base their purchasing decisions on price if all other things are equal. The exact quantity bought for each price level is described in the demand schedule. It's then plotted on a graph to show the demand curve.

The demand curve shows just the relationship between price and quantity. If one of the other determinants changes, the entire demand curve shifts.

If the quantity demanded responds a lot to price, then it's known as elastic demand. If the volume doesn't change much, regardless of price, that's inelastic demand.

Income. When income rises, so will the quantity demanded. When income falls, so will demand. But if your income doubles, you won't always buy twice as much of a particular good or service. There's only so many pints of ice cream you'd want to eat, no matter how wealthy you are. That's where the concept of *marginal utility* comes into the picture. The first pint of ice cream tastes delicious. You might have another. But after that, the marginal utility starts to decrease to the point where you don't want any more.

Prices of related goods or services. The price of complementary goods or services raises the cost of using the product you demand, so you'll want less.

For example, when gas prices rose to \$4 a gallon in 2008, the demand for Hummers fell. Gas is a complementary good to Hummers. The cost of driving a Hummer rose along with gas prices. The opposite reaction occurs when the price of a substitute rises. When that happens, people will want more of the good or service and less of its substitute. That's why Apple continually innovates with its iPhones and iPods. As soon as a substitute, such as a new Android phone, appears at a lower price, Apple comes out with a better product. Then the Android is no longer a substitute.

Tastes. When the public's desires, emotions, or preferences change in favor of a product, so does the quantity demanded. Likewise, when tastes go against it, that depresses the amount demanded. Brand advertising tries to increase the desire for consumer goods. For example, Buick spent millions to make you think its cars are not only for older people.

Expectations. When people expect that the value of something will rise, they demand more of it. That explains the housing asset bubble of 2005. Housing prices rose, but people bought more because they expected the price to continue to go up. Prices increased even more until the bubble burst in 2006. Between 2007 and 2011, housing prices fell 30%. But the quantity demanded didn't grow. Why? People expected prices to continue falling. Record levels of foreclosures entered the market due to the subprime mortgage crisis. Demand didn't increase until people expected future prices would, too.

Number of buyers in the market. The number of consumers affects overall, or "aggregate," demand. As more buyers enter the market, demand rises. That's true even if prices don't change. That was another reason for the housing bubble. Low-cost and subprime mortgages increased the number of people who could afford a house. The total number of buyers in the market expanded. This increased demand for housing. When housing prices started to fall, many realized they couldn't afford their mortgages. At that

point, they foreclosed. That reduced the number of buyers and drove down demand.

Meaning And Determinants Of Demand

We have often heard the phrase ‘there is a huge demand for product XYZ in the market’. But what does this exactly mean? What constitutes the demand for a product in the economy? Let us learn about the concept of demand and the determinants of demand in a market.

What is Demand?

Demand in terms of economics may be explained as the consumers’ willingness and ability to purchase or consume a given item/good. Furthermore, the determinants of demand go a long way in explaining the demand for a particular good.

For instance, an increase in the price of a good will lead to a decrease in the quantity that may be demanded by consumers. Similarly, a decrease in the cost or selling price of a good will most likely lead to an increase in the demanded quantity of the goods.

This indicates the existence of an inverse relationship between the price of the article and the quantity demanded by consumers. This is commonly known as the law of demand and can be graphically represented by a line with a downward slope.

The graphical representation is known as the demand curve. The determinants of demand are factors that cause fluctuations in the economic demand for a product or a service.

Determinants of Demand

Some of the important determinants of demand are as follows,

Price of the Product

People use price as a parameter to make decisions if all other factors remain constant or equal. According to the law of demand, this implies an increase in demand follows a reduction in price and a decrease in demand follows an increase in the price of similar goods.

The demand curve and the demand schedule help determine the demand quantity at a price level. An elastic demand implies a robust change quantity accompanied by a change in price. Similarly, an inelastic demand implies that volume does not change much even when there is a change in price.

Clear Concept of Demand

One of the most important building blocks of economic analysis is the concept of demand. When economists refer to demand, they usually have in mind not just a single quantity demanded, but what is called a demand curve. A demand curve traces the quantity of a good or service that is demanded at successively different prices. The most famous law in economics, and the one that economists are most sure of, is the law of demand. On this law is built almost the whole edifice of economics. The law of demand states that when the price of a good rises, the amount demanded falls, and when the price falls, the amount demanded rises.

Some of the modern evidence for the law of demand is from econometric studies which show that, all other things being equal, when the price of a good rises, the amount of it demanded decreases. How do we know that there are no instances in which the amount demanded rises and the price rises? A few instances have been cited, but they almost always have an explanation that takes into account something other than price. Nobel Laureate George Stigler responded years ago that if any economist found a true counterexample, he would be “assured of immortality, professionally speaking, and rapid promotion.”

And because, wrote Stigler, most economists would like either reward, the fact that no one has come up with an exception to the law of demand shows how rare the exceptions must be. But the reality is that if an economist reported an instance in which consumption of a good rose as its price rose, other economists would assume that some factor other than price caused the increase in demand. The main reason economists believe so strongly in the law of demand is that it is so plausible, even to noneconomists. Indeed, the law of demand is ingrained in our way of thinking about everyday things. Shoppers buy more strawberries when

they are in season and the price is low. This is evidence for the law of demand: only at the lower, in-season price are consumers willing to buy the higher amount available. Similarly, when people learn that frost will strike orange groves in Florida, they know that the price of orange juice will rise.

The price rises in order to reduce the amount demanded to the smaller amount available because of the frost. This is the law of demand. We see the same point every day in countless ways. No one thinks, for example, that the way to sell a house that has been languishing on the market is to raise the asking price. Again, this shows an implicit awareness of the law of demand: the number of potential buyers for any given house varies inversely with the asking price. Indeed, the law of demand is so ingrained in our way of thinking that it is even part of our language. Think of what we mean by the term *on sale*. We do not mean that the seller raised the price. We mean that he or she lowered it. The seller did so in order to increase the amount of goods demanded. Again, the law of demand.

Economists, as is their wont, have struggled to think of exceptions to the law of demand. Marketers have found them. One of the best examples was a new car wax. Economist Thomas Nagle points out that when one particular car wax was introduced, it faced strong resistance until its price was raised from \$.69 to \$1.69. The reason, according to Nagle, was that buyers could not judge the wax's quality before purchasing it. Because the quality of this particular product was so important—a bad product could ruin a car's finish—consumers “played it safe by avoiding cheap products that they believed were more likely to be inferior.”

Many noneconomists are skeptical of the law of demand. A standard example they give of a good whose quantity demanded will not fall when the price increases is water. How, they ask, can people reduce their use of water?

But those who come up with that example think of drinking water, or using it in a household, as the only possible uses. Even for such uses, there is room to reduce consumption when the price of water rises. Households can do larger loads of laundry, or shower instead of bathe, for example. The main users of water,

however, are agriculture and industry. Farmers and manufacturers can substantially alter the amount of water used in production.

Farmers, for example, can do so by changing crops or by changing irrigation methods for given crops. It is not just price that affects the quantity demanded. Income affects it too. As real income rises, people buy more of some goods (which economists call normal goods) and less of what are called inferior goods. Urban mass transit and railroad transportation are classic examples of inferior goods.

That is why the usage of both of these modes of travel declined so dramatically as postwar incomes were rising and more people could afford automobiles. Environmental quality is a normal good, which is a major reason that Americans have become more concerned about the environment in recent decades. Another influence on demand is the price of substitutes. When the price of Toyota Tercels rises, all else being equal, demand for Tercels falls and demand for Nissan Sentras, a substitute, rises. Also important is the price of complements, or goods that are used together. When the price of gasoline rises, the demand for cars falls.

Demand is economic want backed up by purchasing power. i.e., it is the plan, or relationship, expressing different amounts of a product buyers are willing and able to buy at possible prices, assuming all other non-price factors remain the same. For example, a consumer may be willing to purchase 2 lb of potatoes if the price is \$0.75 per lb. However, the same consumer may be willing to purchase only 1 lb if the price is \$1.00 per lb. A demand schedule can be constructed that shows the quantity demanded at each given price. It can be represented on a graph as a line or curve by plotting the quantity demanded at each price. It can also be described mathematically by a demand equation. The main determinants of the quantity one is willing to purchase will typically be the price of the good, one's level of income, personal tastes, the price of substitute goods, and the price of complementary goods.

If the price of widgets is originally \$1.00 and people are buying 100, they may change to 90 for two reasons. One reason

is that the price may rise to \$2.00. The other reason is that one of the factors that is assumed to be constant may change, so that even though the price has not changed, quantity will. Economists distinguish these two cases. In the first case the demand relationship or schedule has not changed, but there has been movement within the relationship. Economists call a change of this sort a change in quantity demanded. The second sort of change is an alteration of the relationship. The original pairing of price and quantity is destroyed and replaced by a new pairing. Economists call this sort of change a change in demand.

It is important to realize that though the demand relationship looks concrete when it is illustrated with a table or graph, in everyday life demand curves are hidden. A demand curve refers to what people would do if various prices were charged, and very rarely are enough prices charged so a clear demand curve can be seen.

This is not to say that the concept is of no importance to people who sell. They may not be interested in the demand curve as a relationship, but they do find it a boundary or constraint on their behaviour. If there were an actual widget seller facing the demand curve in our demand table, he would find that he could not sell more than 90 widgets if he wanted to charge \$2.00. He could of course sell fewer if he wanted to. He could sell only 70 at \$2.00, but if he did this, he would earn far less than he could. If he wanted to sell more than 90, he would have to lower his price.

<i>A Demand Curve</i>	
<i>Price of Widgets</i>	<i>Number of Widgets People Want to Buy</i>
\$1.00	100
\$2.00	90
\$3.00	70
\$4.00	40

Thus, to an actual businessman the demand curve is important as a limitation on what he can do. A businessman may not know

exactly where the demand curve is, and he may not think of it as fixed. Advertising—either informing or persuading people—can move the boundary. As we proceed further, we will see that there are still other ways to view the demand curve in addition to seeing it as a mathematical relationship and as a boundary that limits sellers.

The shape of the aggregated demand curve can be convex or concave, possibly depending on income distribution. The capacity to buy is sometimes used to characterize demand as being merely an alternate form of supply.

Joint Demand

In the process of production, various raw materials enter into the finished product. These different raw materials are said to be jointly demanded as the demand for each is derived from the ultimate demand for the final commodity into the production of which they enter. Thus coal and iron ore are jointly demanded in order to produce steel.

From what has been said, it is evident that joint demand is practically universal in the sphere of the production of final products (i.e. those for final consumption). With regard to the relation between joint demand and price, if for any reason there is an increase in the demand for the ultimate product, then this will be reflected in the demand for the constituent products and their prices. As a result, that constituent which has a relatively inelastic supply will increase most in price and it will thus absorb a greater proportion of the increased price of the ultimate product. But it must be stated in conclusion that the conditions of joint demand will have no lasting effect on price.

Joint Supply

It is invariably true that in the course of most production processes, various by-products emerge together with the ultimate product. In other words, we have a case of joint supply, i.e. two or more goods being derived from a common source. Even in the extractive industries we have beef and hides from the rearing of cattle; mutton and wool from sheep; cotton-lint and cotton-seeds

from the cotton plant. In examining the relation between joint supply and price we have first to recognize two important facts, namely:

- That the commodities derived from a common origin are forthcoming in what are more or less unvarying proportions to one another.
- There is no similar connection between the separate demands for each of the commodities jointly produced.

It follows, therefore, that the total of the prices obtained for the individual joint products must be sufficient to cover their joint expenses of production, but the relative prices of the individual products are more difficult to assess, and lie outside the scope of this elementary discussion.

Composite Demand

It is a characteristic fact to-day that one commodity has more than one use in production, and is therefore capable of being used in the manufacture of entirely different products. For example, leather is demanded for the manufacture of boots and shoes, handbags, travelling-trunks, and fancy leather goods. Again, steel is used in many industries such as building construction, engineering, ship-building. The demands for leather or steel in the cases already mentioned are competitive demands relative to one another, and it is this fact that gives us the clue to the problem of price determination in the case of products the demand for which is composite. If, for any reason, there is an increased demand for the commodity in one of its uses, e.g. boots and shoes, the immediate effect is to intensify the scarcity of the commodity in relation to its alternative uses and raise its price, although not, of course, equally.

The extent of this increase in price will depend on the *elasticity of the supply* of the commodity in question. Thus in the case of steel, new supplies will be forthcoming fairly readily, but in the case of leather, depending ultimately on the supply of hides, new supplies will be put on the market only after an appreciable lapse of time, and with difficulty. Where new supplies are easily forthcoming the rise in price will not be relatively great nor

prolonged over a period of time. The more difficult it is to increase the supplies of the commodity in question, the greater will be the increase in price and the longer will be the duration of the increase.

Composite Supply

Composite supply exists when there are two or more products which satisfy the same need, e.g. tea, coffee, and cocoa: butter and margarine. In these examples it must be pointed out that no product is in reality a perfect substitute for another. If the supply of one commodity in a composite supply is increased, conditions of demand remaining the same, the price for it will fall, and at this lower price there will be a tendency for a greater quantity of it to be taken in the place of the other commodity or commodities, the supply of which for obvious reasons will now tend to diminish.

Change In Demand

A change in demand describes a shift in consumer desire to purchase a particular good or service, irrespective of a variation in its price. The change could be triggered by a shift in income levels, consumer tastes, or a different price being charged for a related product.

Understanding Change In Demand

Demand is an economic principle referring to a consumer's desire to buy things. There are a number of factors that influence market demand for a particularly good or service. The main determinants are:

- Income: How much consumers have to spend.
- Consumer preferences: What types of products are popular at any given moment.
- Buyer expectations: Does the consumer expect the price to rise in the future, perhaps due to limited supply?
- Price: How much does the good or service cost?
- Prices of related items: Are there any substitute goods or services of similar value that cost a lot less?

A change in demand occurs when appetite for goods and services shifts, even though prices remain constant. When the

economy is flourishing and incomes are rising, consumers could feasibly purchase more of everything. Prices will remain the same, at least in the short-term, while the quantity sold increases.

In contrast, demand could be expected to drop at every price during a recession. When economic growth abates, jobs tend to get cut, incomes fall, and people get nervous, refraining from making discretionary expenses and only buying essentials.

Recording Change in Demand

An increase and decrease in total market demand is illustrated in the demand curve, a graphical representation of the relationship between the price of a good or service and the quantity demanded for a given period of time. Typically, the price will appear on the left vertical y-axis, while the quantity demanded is shown on the horizontal x-axis.

The supply and demand curves form an X on the graph, with supply pointing upward and demand pointing downward. Drawing straight lines from the intersection of these two curves to the x- and y-axes yields price and quantity levels based on current supply and demand.

Consequently, a positive change in demand amid constant supply shifts the demand curve to the right, the result being an increase in price and quantity. Alternatively, a negative change in demand shifts the curve left, leading price and quantity to both fall.

Change in Demand vs. Quantity Demanded

It is important not to confuse change in demand with quantity demanded. Quantity demanded describes the total amount of goods or services demanded at any given point in time, depending on the price being charged for them in the marketplace. Change in demand, on the other hand, focuses on all determinants of demand other than price changes.

Example of Change of Demand

When an item becomes fashionable, perhaps due to smart advertising, consumers clamor to buy it. For instance, Apple Inc.'s

iPhone sales have remained fairly constant, despite undergoing various price increases over the years, as many consumers view it as the number one smartphone in the market and are locked into Apple's ecosystem. In various parts of the world, the Apple iPhone has also become a status symbol, illustrating inelastic demand just as Nokia Corp.'s cellphones did in the early 2000s. Technological advancements and fashion trends aren't the only factors that can trigger a change in demand. For example, during the mad cow disease scare, consumers started buying chicken rather than beef, even though the latter's price had not changed.

Chicken could also find itself in favor if the price of another competing poultry products rises significantly. In such a scenario, demand for chicken rockets, despite still costing the same at the supermarket. Alternatively, if there is a perceived increase in the price of gasoline, then there could feasibly be a decrease in the demand for gas-guzzling SUVs, *ceteris paribus*.

CONCEPT OF DEMAND ESTIMATION

Empirical Estimation

Demand and supply relations in a market can be statistically estimated from price, quantity, and other data with sufficient information in the model. This can be done with *simultaneous-equation methods of estimation* in econometrics. Such methods allow solving for the model-relevant "structural coefficients," the estimated algebraic counterparts of the theory.

The *Parameter identification problem* is a common issue in "structural estimation." Typically, data on exogenous variables (that is, variables other than price and quantity, both of which are endogenous variables) are needed to perform such an estimation. An alternative to "structural estimation" is reduced-form estimation, which regresses each of the endogenous variables on the respective exogenous variables.

Macroeconomic uses of demand and supply

Demand and supply have also been generalized to explain macroeconomic variables in a market economy, including the quantity of total output and the general price level. The Aggregate

Demand-Aggregate Supply model may be the most direct application of supply and demand to macroeconomics, but other macroeconomic models also use supply and demand.

Compared to micro-economic uses of demand and supply, different (and more controversial) theoretical considerations apply to such macroeconomic counterparts as *aggregate demand* and *aggregate supply*. Demand and supply may also be used in macroeconomic theory to relate money supply to demand and interest rates.

Demand Shortfalls

A demand shortfall results from the actual demand for a given product being lower than the projected, or estimated, demand for that product. Demand shortfalls are caused by demand overestimation in the planning of new products. Demand overestimation is caused by optimism bias and/or strategic misrepresentation.

The power of supply and demand was understood to some extent by several early Muslim economists, such as Ibn Taymiyyah who illustrates: "If desire for goods increases while its availability decreases, its price rises. On the other hand, if availability of the good increases and the desire for it decreases, the price comes down."

The phrase "supply and demand" was first used by James Denham-Steuart in his *Inquiry into the Principles of Political Economy*, published in 1767. Adam Smith used the phrase in his 1776 book *The Wealth of Nations*, and David Ricardo titled one chapter of his 1817 work *Principles of Political Economy and Taxation* "On the Influence of Demand and Supply on Price". In *The Wealth of Nations*, Smith generally assumed that the supply price was fixed but that its "merit" (value) would decrease as its "scarcity" increased, in effect what was later called the law of demand. Ricardo, in *Principles of Political Economy and Taxation*, more rigorously laid down the idea of the assumptions that were used to build his ideas of supply and demand. Antoine Augustin Cournot first developed a mathematical model of supply and demand in his 1838 *Researches on the Mathematical Principles of the Theory of Wealth*.

During the late 19th century the marginalist school of thought emerged. This field mainly was started by Stanley Jevons, Carl Menger, and Leon Walras. The key idea was that the price was set by the most expensive price, that is, the price at the margin. This was a substantial change from Adam Smith's thoughts on determining the supply price.

In his 1870 essay "On the Graphical Representation of Supply and Demand", Fleeming Jenkin drew for the first time the popular graphic of supply and demand which, through Marshall, eventually would turn into the most famous graphic in economics. The model was further developed and popularized by Alfred Marshall in the 1890 textbook *Principles of Economics*. Along with Leon Walras, Marshall looked at the equilibrium point where the two curves crossed. They also began looking at the effect of markets on each other.

Criticism

At least two assumptions are necessary for the validity of the standard model: first, that supply and demand are independent; and second, that supply is "constrained by a fixed resource"; If these conditions do not hold, then the Marshallian model cannot be sustained. Sraffa's critique focused on the inconsistency (except in implausible circumstances) of partial equilibrium analysis and the rationale for the upward-slope of the supply curve in a market for a produced consumption good. The notability of Sraffa's critique is also demonstrated by Paul A. Samuelson's comments and engagements with it over many years, for example: "What a cleaned-up version of Sraffa (1926) establishes is how *nearly empty* are *all* of Marshall's partial equilibrium boxes. To a logical purist of Wittgenstein and Sraffa class, the *Marshallian partial equilibrium* box of *constant* cost is even more empty than the box of *increasing* cost."

Aggregate excess demand in a market is the difference between the quantity demanded and the quantity supplied as a function of price. In the model with an upward-sloping supply curve and downward-sloping demand curve, the aggregate excess demand function only intersects the axis at one point, namely, at the point

where the supply and demand curves intersect. The Sonnenschein-Mantel-Debreu theorem shows that the standard model cannot be rigorously derived in general from the theory of general equilibrium.

The model of prices being determined by supply and demand assume perfect competition. But: "economists have no adequate model of how individuals and firms adjust prices in a competitive model. If all participants are price-takers by definition, then the actor who adjusts prices to eliminate excess demand is not specified"

DEMAND FORECASTING

Demand Forecasting is the process in which historical sales data is used to develop an estimate of an expected forecast of customer demand. To businesses, Demand Forecasting provides an estimate of the amount of goods and services that its customers will purchase in the foreseeable future. Critical business assumptions like turnover, profit margins, cash flow, capital expenditure, risk assessment and mitigation plans, capacity planning, etc. are dependent on Demand Forecasting.

Demand

Supply and demand is an economic model of price determination in a market. It concludes that in a competitive market, price will function to equalize the quantity demanded by consumers, and the quantity supplied by producers, resulting in an economic equilibrium of price and quantity.

The Graphical Representation of Supply and Demand

The supply-demand model is a partial equilibrium model representing the determination of the price of a particular good and the quantity of that good which is traded. Although it is normal to regard the quantity demanded and the quantity supplied as functions of the price of the good, the standard graphical representation, usually attributed to Alfred Marshall, has price on the vertical axis and quantity on the horizontal axis, the opposite

of the standard convention for the representation of a mathematical function.

Determinants of supply and demand other than the price of the good in question, such as consumers' income, input prices and so on, are not explicitly represented in the supply-demand diagram. Changes in the values of these variables are represented by shifts in the supply and demand curves. By contrast, responses to changes in the price of the good are represented as movements along unchanged supply and demand curves.

Supply Schedule

The supply schedule, depicted graphically as the supply curve, represents the amount of some good that producers are willing and able to sell at various prices, assuming *ceteris paribus*, that is, assuming all determinants of supply other than the price of the good in question, such as technology and the prices of factors of production, remain the same.

Under the assumption of perfect competition, supply is determined by marginal cost. Firms will produce additional output as long as the cost of producing an extra unit of output is less than the price they will receive.

Demand Schedule

The demand schedule, depicted graphically as the demand curve, represents the amount of some good that buyers are willing and able to purchase at various prices, assuming all determinants of demand other than the price of the good in question, such as income, personal tastes, the price of substitute goods, and the price of complementary goods, remain the same. Following the law of demand, the demand curve is almost always represented as downward-sloping, meaning that as price decreases, consumers will buy more of the good.

Just as the supply curves reflect marginal cost curves, demand curves are determined by marginal utility curves. Consumers will be willing to buy a given quantity of a good, at a given price, if the marginal utility of additional consumption is equal to the opportunity cost determined by the price, that is, the marginal

utility of alternative consumption choices. The demand schedule is defined as the *willingness* and *ability* of a consumer to purchase a given product in a given frame of time. As described above, the demand curve is generally downward-sloping. There may be rare examples of goods that have upward-sloping demand curves. Two different hypothetical types of goods with upward-sloping demand curves are Giffen goods (an inferior but staple good) and Veblen goods (goods made more fashionable by a higher price).

Forecasting

Forecasting sales and profits, particularly on a short-term basis (one year to three years), is essential to planning for business success. This process, estimating future business performance based on the actual results from prior periods, enables the business owner/manager to modify the operation of the business on a timely basis.

This allows the business to avoid losses or major financial problems should some future results from operations not conform with reasonable expectations. Forecasts-or Pro Forma Income Statements and Cash Flow Statements as they are usually called-also provide the most persuasive management tools to apply for loans or attract investor money. As a business expands, there will inevitably be a need for more money than can be internally generated from profits. Next, let's examine some facts affecting Forecasting sales and profits.

Facts Affecting Pro Forma Statements

Preparation of Forecasts (Pro Forma Statements) requires assembling a wide array of pertinent, verifiable facts affecting your business and its past performance. These include:

Data from prior financial statements, particularly:

- a. Previous sales levels and trends.
- b. Past gross percentages.
- c. Average past general, administrative, and selling expenses necessary to generate your former sales volumes.
- d. Trends in the company's need to borrow (supplier, trade credit, and bank credit) to support various levels of

inventory and trends in accounts receivable required to achieve previous sales volumes.

Unique company data, particularly:

- a. Plant capacity.
- b. Competition.
- c. Financial constraints.
- d. Personnel availability.

Industry-wide factors, including:

- a. Overall state of the economy.
- b. Economic status of your industry within the economy.
- c. Population growth.
- d. Elasticity of demand for the product or service your business provides (Demand is said to be “elastic” if it decreases as prices increase, a demonstration that consumers can do without or with less of the goods or service. If demand for something is relatively steady as prices increase, it is “inelastic”).
- e. Availability of raw materials.

Once these factors are identified, they may be used in Pro Formas, which estimate the level of sales, expense, and profitability that seem possible in a future period of operations.

The Pro Forma Income Statement

In preparing the Pro Forma Income Statement, the estimate of total sales during a selected period is the most critical “guesstimate:” Employ business experience from past financial statements. Get help from management and sales people in developing this all-important number.

Then assume, for example, that a 10 percent increase in sales volume is a realistic and attainable goal. Multiply last year’s net sales by 1.10 to get this year’s estimate of total net sales. Next, break down this total, month by month, by looking at the historical monthly sales volume. From this you can determine what percentage of total annual sales fell on the average in each of those months over a minimum of the past three years. You may find that

75 percent of total annual sales volume was realized during the six months from July through December in each of those years and that the remaining 25 percent of sales was spread fairly evenly over the first six months of the year. Next, estimate the cost of goods sold by analysing operating data to determine on a monthly basis what percentage of sales has gone into cost of goods sold in the past.

This percentage can then be adjusted for expected variations in costs, price trends, and efficiency of operations. Operating expenses (sales, general and administrative expenses, depreciation, and interest), other expenses, other income, and taxes can then be estimated through detailed analysis and adjustment of what they were in the past and what you expect them to be in the future.

Comparison with Actual Monthly Performance

Putting together this information month by month for a year into the future will result in your business's Pro Forma Statement of Income.

Use it to compare with the actual monthly results from operations. Preparation of the information is summarized below:

Revenue (Sales)

- List the departments within the business. For example, if your business is appliance sales and service, the departments would include new appliances, used appliances, parts, in-shop service, on-site service.
- In the "Estimate" columns, enter a reasonable projection of monthly sales for each department of the business. Include cash and on-account sales. In the "Actual" columns, enter the actual sales for the month as they become available.
- Exclude from the Revenue section any revenue not strictly related to the business.

Cost of Sales

- Cite costs by department of the business, as above.

- In the “Estimate” columns, enter the cost of sales estimated for each month for each department. For product inventory, calculate the cost of the goods sold for each department (beginning inventory plus purchases and transportation costs during the month minus the inventory). Enter “Actual” costs each month as they accrue.

Gross Profit

Subtract the total cost of sales from the total revenue.

Expenses

- *Salary Expenses*: Base pay plus overtime.
- *Payroll Expenses*: Include paid vacations, sick leave, health insurance, unemployment insurance, Social Security taxes.
- *Outside Services*: Include costs of subcontracts, overflow work farmed-out, special or one-time services.
- *Supplies*: Services and items purchased for use in the business, not for resale.
- *Repairs and Maintenance*: Regular maintenance and repair, including periodic large expenditures, such as painting or decorating.
- *Advertising*: Include desired sales volume, classified directory listing expense, etc.
- *Car, Delivery and Travel*: Include charges if personal car is used in the business. Include parking, tolls, mileage on buying trips, repairs, etc.
- *Accounting and Legal*: Outside professional services.
- *Rent*: List only real estate used in the business.
- Telephone.
- *Utilities*: Water, heat, light, etc.
- *Insurance*: Fire or liability on property or products, worker’s compensation.
- *Taxes*: Inventory, sales, excise, real estate, others.
- Interest.

- *Depreciation*: Amortization of capital assets.
- *Other Expenses (specify each)*: Tools, leased equipment, etc.
- *Miscellaneous (unspecified)*: Small expenditures without separate accounts.

Net Profit

- To find net profit, subtract total expenses from gross profit.

The Pro Forma Statement of Income, prepared on a monthly basis and culminating in an annual projection for the next business fiscal year, should be revised not less than quarterly.

It must reflect the actual performance achieved in the immediately preceding three months to ensure its continuing usefulness as one of the two most valuable planning tools available to management. Should the Pro Forma reveal that the business will likely not generate a profit from operations, plans must immediately be developed to identify what to do to at least break even –increase volume, decrease expenses, or put more owner capital in to pay some debts and reduce interest expenses.

Break-Even Analysis

“Break-Even” means a level of operations at which a business neither makes a profit nor sustains a loss. At this point, revenue is just enough to cover expenses. Break-Even Analysis enables you to study the relationship of volume, costs, and revenue. Break-Even requires the business owner/manager to define a sales level—either in terms of revenue dollars to be earned or in units to be sold within a given accounting period—at which the business would earn a before tax net profit of zero. This may be done by employing one of various formula calculations to the business estimated sales volume, estimated fixed costs, and estimated variable costs. Generally, the volume and cost estimates assume the following conditions:

- A change in sales volume will not affect the selling price per unit;
- Fixed expenses (rent, salaries, administrative and office expenses, interest, and depreciation) will remain the same at all volume levels; and

- Variable expenses (cost of goods sold, variable labour costs, including overtime wages and sales commissions) will increase or decrease in direct proportion to any increase or decrease in sales volume.

Two methods are generally employed in Break-Even Analysis, depending on whether the break-even point is calculated in terms of sales dollar volume or in number of units that must be sold.

Break-Even Point in Sales Dollars

The steps for calculating the first method are shown below:

- Obtain a list of expenses incurred by the company during its past fiscal year.
- Separate the expenses listed in Step 1 into either a variable or a fixed expense classification.
- Express the variable expenses as a percentage of sales. In the condensed income statement of the Small Business Specialities Co. (below), net sales were \$1,200,000. In Step 2, variable expenses were found to amount to \$720,000. Therefore, variable expenses are 60 percent of net sales (\$720,000 divided by \$1,200,000). This means that 60 cents of every sales dollar is required to cover variable expenses. Only the remainder, 40 cents of every dollar, is available for fixed expenses and profit.
- Substitute the information gathered in the preceding steps in the following basic break-even formula to calculate the break-even point.

Remember: Increased sales do not necessarily mean increased profits. If you know your company's break-even point, you will know how to price your product to make a profit. If you cannot make an acceptable profit, alter or sell your business before you lose your retained earnings.

Demand Forecasting types

Demand Forecasting can be broadly classified based on the level of detailing, time span considered and the scope of market considered.

Outlined below are the major types of Demand Forecasting:

- **Passive Demand Forecasting:** Passive Demand Forecasting is carried out for stable businesses with very conservative growth plans. Simple extrapolations of historical data is carried out with minimal assumptions. This is a rare type of forecasting limited to small and local businesses.
- **Active Demand Forecasting:** Active Demand Forecasting is carried out for scaling and diversifying businesses with aggressive growth plans in terms of marketing activities, product portfolio expansion and consideration of competitor activities and external economic environment.
- **Short-term Demand Forecasting:** Short-term Demand Forecasting is carried out for a shorter term period of 3 months to 12 months. In the short term, the seasonal pattern of demand and the effect of tactical decisions on the customer demand are taken into consideration.
- **Medium to long-term Demand Forecasting:** Medium to long-term Demand Forecasting is typically carried out for more than 12 months to 24 months in advance (36-48 months in certain businesses). Long-term Forecasting drives the business strategy planning, sales and marketing planning, financial planning, capacity planning, capital expenditure, etc.
- **External macro level Demand Forecasting:** This type of Forecasting deals with the broader market movements which depend on the macroeconomic environment. External Forecasting is carried out for evaluating the strategic objectives of a business like product portfolio expansion, entering new customer segments, technological disruptions, a paradigm shift in consumer behavior and risk mitigation strategies.
- **Internal business level Demand Forecasting:** As the name suggests, this type of Forecasting deals with internal operations of the business such as product category, sales division, financial division, and manufacturing group. This

includes annual sales forecast, estimation of COGS, net profit margin, cash flow, etc.

Demand Forecasting examples

Some real-world practical examples of Demand Forecasting are – A leading car maker, refers to the last 12 months of actual sales of its cars at model, engine type, and color level; and based on the expected growth, forecasts the short-term demand for the next 12 month for purchase, production and inventory planning purposes. A leading food manufacturing company refers to the last 24 months of actual sales of its highly seasonal products like soups and mashed potatoes. An analysis is carried out at the flavor and packaging size level.

Then based on the market potential, demand is forecasted for the next 12 to 24 months for sourcing of key ingredients like tomatoes, potatoes, etc. and for capacity planning and evaluating the need for external co-packing.

Importance of Demand Forecasting

Demand Forecasting is the pivotal business process around which strategic and operational plans of a company are devised. Based on the Demand Forecast, strategic and long-range plans of a business like budgeting, financial planning, sales and marketing plans, capacity planning, risk assessment and mitigation plans are formulated.

Short to medium term tactical plans like pre-building, make-to-stock, make-to-order, contract manufacturing, supply planning, network balancing, etc. are execution based. Demand Forecasting also facilitates important management activities like decision making, performance evaluation, judicious allocation of resources in a constrained environment and business expansion planning.

Techniques of Demand Forecasting

Broadly speaking, there are two approaches to demand forecasting—one is to obtain information about the likely purchase behaviour of the buyer through collecting expert's opinion or by conducting interviews with consumers, the other is to use past experience as a guide through a set of statistical techniques. Both

these methods rely on varying degrees of judgment. The first method is usually found suitable for short-term forecasting, the latter for long-term forecasting. There are specific techniques which fall under each of these broad methods.

Simple Survey Method

For forecasting the demand for existing product, such survey methods are often employed. In this set of methods, we may undertake the following exercise.

- 1) Experts Opinion Poll: In this method, the experts on the particular product whose demand is under study are requested to give their 'opinion' or 'feel' about the product. These experts, dealing in the same or similar product, are able to predict the likely sales of a given product in future periods under different conditions based on their experience. If the number of such experts is large and their experience-based reactions are different, then an average-simple or weighted -is found to lead to unique forecasts. Sometimes this method is also called the 'hunch method' but it replaces analysis by opinions and it can thus turn out to be highly subjective in nature.
- 2) Reasoned Opinion-Delphi Technique: This is a variant of the opinion poll method. Here is an attempt to arrive at a consensus in an uncertain area by questioning a group of experts repeatedly until the responses appear to converge along a single line. The participants are supplied with responses to previous questions (including seasonings from others in the group by a coordinator or a leader or operator of some sort). Such feedback may result in an expert revising his earlier opinion. This may lead to a narrowing down of the divergent views (of the experts) expressed earlier. The Delphi Techniques, followed by the Greeks earlier, thus generates "reasoned opinion" in place of "unstructured opinion"; but this is still a poor proxy for market behaviour of economic variables.
- 3) Consumers Survey-Complete Enumeration Method: Under this, the forecaster undertakes a complete survey of all consumers whose demand he intends to forecast,

Once this information is collected, the sales forecasts are obtained by simply adding the probable demands of all consumers. The principle merit of this method is that the forecaster does not introduce any bias or value judgment of his own. He simply records the data and aggregates. But it is a very tedious and cumbersome process; it is not feasible where a large number of consumers are involved. Moreover if the data are wrongly recorded, this method will be totally useless.

- 4) **Consumer Survey-Sample Survey Method:** Under this method, the forecaster selects a few consuming units out of the relevant population and then collects data on their probable demands for the product during the forecast period. The total demand of sample units is finally blown up to generate the total demand forecast. Compared to the former survey, this method is less tedious and less costly, and subject to less data error; but the choice of sample is very critical. If the sample is properly chosen, then it will yield dependable results; otherwise there may be sampling error. The sampling error can decrease with every increase in sample size
- 5) **End-user Method of Consumers Survey:** Under this method, the sales of a product are projected through a survey of its end-users. A product is used for final consumption or as an intermediate product in the production of other goods in the domestic market, or it may be exported as well as imported. The demands for final consumption and exports net of imports are estimated through some other forecasting method, and its demand for intermediate use is estimated through a survey of its user industries.

ELASTICITY OF DEMAND: MEANING MEASUREMENT, DETERMINING FACTORS AND IMPORTANCE

Elasticity

An important concept in understanding supply and demand theory is elasticity. In this context, it refers to how supply and

demand change in response to various stimuli. One way of defining elasticity is the percentage change in one variable divided by the percentage change in another variable (known as arch elasticity because it calculates the elasticity over a range of values, in contrast with point elasticity that uses differential calculus to determine the elasticity at a specific point). Thus it is a measure of relative changes.

Often, it is useful to know how the quantity supplied or demanded will change when the price changes. This is known as the price elasticity of demand and the price elasticity of supply. If a monopolist decides to increase the price of their product, how will this affect their sales revenue? Will the increased unit price offset the likely decrease in sales volume? If a government imposes a tax on a good, thereby increasing the effective price, how will this affect the quantity demanded?

If you do not wish to calculate elasticity, a simpler technique is to look at the slope of the curve. Unfortunately, this has units of measurement of quantity over monetary unit (for example, liters per euro, or battleships per million yen), which is not a convenient measure to use for most purposes.

So, for example, if you wanted to compare the effect of a price change of gasoline in Europe versus the United States, there is a complicated conversion between gallons per dollar and liters per euro.

This is one of the reasons why economists often use relative changes in percentages, or elasticity. Another reason is that elasticity is more than just the slope of the function: It is the slope of a function in a coordinate space, that is, a line with a constant slope will have different elasticity at various points.

Let's do an example calculation. We have said that one way of calculating elasticity is the percentage change in quantity over the percentage change in price. So, if the price moves from \$1.00 to \$1.05, and the quantity supplied goes from 100 pens to 102 pens, the slope is $2/0.05$ or 40 pens per dollar. Since the elasticity depends on the percentages, the quantity of pens increased by 2%, and the price increased by 5%, so the price elasticity of supply is $2/5$ or 0.4.

Since the changes are in percentages, changing the unit of measurement or the currency will not affect the elasticity. If the quantity demanded or supplied changes a lot when the price changes a little, it is said to be elastic. If the quantity changes little when the prices changes a lot, it is said to be inelastic. An example of perfectly inelastic supply, or zero elasticity, is represented as a vertical supply curve. Elasticity in relation to variables other than price can also be considered. One of the most common to consider is income. How would the demand for a good change if income increased or decreased? This is known as the income elasticity of demand. For example, how much would the demand for a luxury car increase if average income increased by 10%? If it is positive, this increase in demand would be represented on a graph by a positive shift in the demand curve, because at all price levels, a greater quantity of luxury cars would be demanded.

Another elasticity that is sometimes considered is the cross elasticity of demand, which measures the responsiveness of the quantity demanded of a good to a change in the price of another good. This is often considered when looking at the relative changes in demand when studying complement and substitute goods. Complement goods are goods that are typically utilized together, where if one is consumed, usually the other is also. Substitute goods are those where one can be substituted for the other, and if the price of one good rises, one may purchase less of it and instead purchase its substitute.

Cross elasticity of demand is measured as the percentage change in demand for the first good that occurs in response to a percentage change in price of the second good. For an example with a complement good, if, in response to a 10% increase in the price of fuel, the quantity of new cars demanded decreased by 20%, the cross elasticity of demand would be $-20\%/10\%$ or, -2 .

Price Elasticity

Businesses know that they face demand curves, but rarely do they know what these curves look like. Yet sometimes a business needs to have a good idea of what part of a demand curve looks like if it is to make good decisions. If Rick's Pizza raises its prices

by ten percent, what will happen to its revenues? The answer depends on how consumers will respond. Will they cut back purchases a little or a lot? This question of how responsive consumers are to price changes involves the economic concept of elasticity. Elasticity is a measure of responsiveness. Two words are important here. The word “measure” means that elasticity results are reported as numbers, or elasticity coefficients. The word “responsiveness” means that there is a stimulus-reaction involved. Some change or stimulus causes people to react by changing their behaviour, and elasticity measures the extent to which people react. The most common elasticity measurement is that of price elasticity of demand. It measures how much consumers respond in their buying decisions to a change in price. The basic formula used to determine price elasticity is

$$e = (\text{percentage change in quantity}) / (\text{percentage change in price}).$$

(Read that as elasticity is the percentage change in quantity divided by the percentage change in price.)

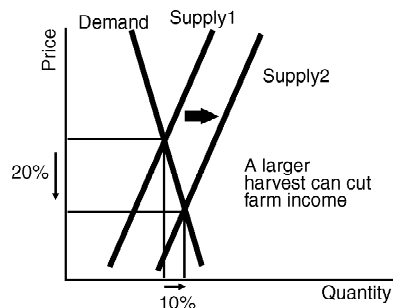
If price increases by 10% and consumers respond by decreasing purchases by 20%, the equation computes the elasticity coefficient as -2. The result is negative because an increase in price (a positive number) leads to a decrease in purchases (a negative number). Because the law of demand says it will always be negative, many economists ignore the negative sign, as we will in the following discussion.

An elasticity coefficient of 2 shows that consumers respond a great deal to a change in price. If, on the other hand, a 10% change in price causes only a 5% change in sales, the elasticity coefficient will be only 1/2. Economists would say in this case that demand is inelastic. Demand is inelastic whenever the elasticity coefficient is less than one. When it is greater than one, economists say that demand is elastic.

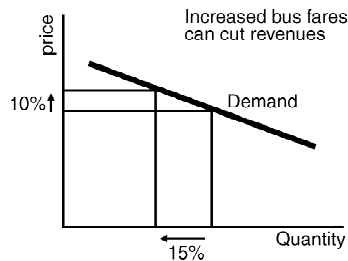
Products that have few good substitutes generally have a lower elasticity of demand than products with many substitutes. As a result, more broadly defined products have a lower elasticity than narrowly defined products. The price elasticity of demand

for meat will be lower than the price elasticity of pork, and the price elasticity for soft drinks will be less elastic than the price elasticity for colas, which in turn will be less elastic than the price elasticity for Pepsi.

Time plays an important role in determining both consumer and producer responsiveness for many items. The longer people have to make adjustments, the more adjustments they will make. When the price of gasoline rose rapidly in the late 1970s as a result of the OPEC cartel, the only adjustment consumers could initially make was to drive less. With time, they could also move closer to work or find jobs closer to home, and switch to more fuel-efficient cars. The concept of elasticity can help explain some situations that at first glance may seem puzzling. If American farmers all have excellent harvests, they may have a very poor year financially. They may be better off if they all have mediocre harvests. If a bus company decides it needs more revenue and tries to get it by raising fares, its revenues may decrease rather than increase.



In the case of the farmers, the key to their problem is that the demand curve for their products is quite inelastic. This means that if the harvest is unusually good, a large drop in price is necessary to encourage consumers to use the additional grain. If the elasticity coefficient is .5, for example, and the harvest is 10% larger than the previous year, then a 20% drop in prices will occur (assuming that the many things that we keep constant in drawing the demand curve have remained constant). Because this price reduction more than offsets the effect of the larger harvest, the average farmer's income drops.



For the bus company, the key is that demand is elastic. For example, suppose that the elasticity is 1.5. Then, if price is raised by 10%, quantity (ridership) must drop by 15%. But the drop in ridership more than offsets the increase in price, and so revenue will drop.

Just as we can measure how responsive buyers are to a change in price, we can measure how responsive sellers are. This measurement, the price elasticity of supply, has the same formula as price elasticity of demand, only the quantity in the formula will refer to the quantity that sellers will sell.

As with demand elasticity, supply elasticity depends on the amount of time available for adjustment. In the very short run, there may be no adjustments sellers can make, which would mean a perfectly vertical supply curve. For example, if on December 1 the price of apples doubles, there will be minimal effect on the number of apples available to the consumer. Producers cannot make adjustments until a new growing season begins. In the short run producers can use their facilities more or less intensively. In the apple example, they can vary the amounts of pesticides and the amount of Labour they use to pick the apples. Finally, in the long run, not only can producers change their facilities, but they can also leave the industry or new producers may enter it. In our apple example, new orchards can be planted or old ones destroyed.

Price Elasticity of Demand

In economics, the price elasticity of demand (PED) is an elasticity that measures the nature and degree of the relationship between changes in quantity demanded of a good and changes in its price. Price elasticity of demand is an

elasticity that measures the nature and degree of the relationship between changes in quantity demanded of a good and changes in its price. For example, if, in response to a 10 % fall in the price of a good, the quantity demanded increases by 20 %, the price elasticity of demand would be $20 \% / (-10 \%) = -2$.

In general, a fall in the price of a good is expected to increase the quantity demanded, so the price elasticity of demand is negative as above. Note that in economics literature the minus sign is often omitted and the elasticity is given as an absolute value. Because both the denominator and numerator of the fraction are percent changes, price elasticities of demand are dimensionless numbers and can be compared even if the original calculations were performed using different currencies or goods.

An example of a good with a highly inelastic demand curve is salt: people need salt, so for even relatively large changes in the price of salt, the amount demanded will not be significantly altered. Similarly, a product with a highly elastic demand curve is red cars: if the price of red cars went up even a small amount, demand is likely to go down since substitutes are readily available for purchase (cars of other colours).

It may be possible that quantity demanded for a good rises as its price rises, even under conventional economic assumptions of consumer rationality. Two such classes of goods are known as Giffen goods or Veblen goods. Another case is the price inflation during an economic bubble. Various research methods are used to calculate price elasticity:

- Test markets
- Analysis of historical sales data
- Conjoint analysis.

Mathematical Definition

The formula used to calculate the coefficient of price elasticity of demand is

$$Ed = \left| \frac{\% \text{ change in quantity demanded of product X}}{\% \text{ chan in price of product X}} \right| = \frac{\Delta Q_d / Q_d}{\Delta P_d / P_d}$$

Or, using the differential calculus:

$$Ed = \frac{P}{Q} \times \frac{\partial Q}{\partial P}$$

where:

P = price

Q = quantity + $x(1/2) - x$

Elasticity and Revenue

When the price elasticity of demand for a good is elastic ($|Ed| > 1$), the percentage change in quantity is greater than that in price. Hence, when the price is raised, the total revenue of producers falls, and vice versa. When the price elasticity of demand for a good is inelastic ($|Ed| < 1$), the percentage change in quantity is smaller than that in price. Hence, when the price is raised, the total revenue of producers rises, and vice versa.

When the price elasticity of demand for a good is unit elastic (or unitary elastic) ($|Ed| = 1$), the percentage change in quantity is equal to that in price. Hence, when the price is raised, the total revenue remains unchanged. The demand curve is a rectangular hyperbola. When the price elasticity of demand for a good is perfectly elastic (Ed is undefined), any increase in the price, no matter how small, will cause demand for the good to drop to zero. Hence, when the price is raised, the total revenue of producers falls to zero. The demand curve is a horizontal straight line. A ten-dollar banknote is an example of a perfectly elastic good; nobody would pay \$10.01, yet everyone will pay \$9.99 for it.

When the price elasticity of demand for a good is perfectly inelastic ($Ed = 0$), changes in the price do not affect the quantity demanded for the good. The demand curve is a vertical straight line; this violates the law of demand. An example of a perfectly inelastic good is a human heart for someone who needs a transplant; nobody would buy more than the exact amount of hearts demanded, no matter how low the price is.

Point-price Elasticity

- Point Elasticity = (% change in Quantity)/(% change in Price)

- Point Elasticity = $(\Delta/Q)/(\Delta/P)$
- Point Elasticity = $(P \Delta Q)/(Q \Delta P)$
- Point Elasticity = $(P/Q)(\Delta Q/\Delta P)$ Note: In the limit (or “at the margin”), “ $(\Delta Q/\Delta P)$ ” is the derivative of the demand function with respect to P. “Q” means ‘quantity’ and “P” means ‘price.’

Example. demand curve: $Q = 1,000 - .6P$ a.) Given this demand curve determine the point price elasticity of demand at $P = 80$ and $P = 40$ as follows. i.) obtain the derivative of the demand function

when it's expressed Q as a function of P. $\frac{\partial Q}{\partial P} = -.6$ ii.) next apply the above equation to the sought ordered pairs: (40, 976), (80, 952)

$$E_p = \frac{\partial Q}{\partial P} = \frac{P}{Q} \quad e = -.6(40/976) = -.02 \quad e = -.6(80/952) = -.05$$

Price Elasticity of Supply

In economics, the price elasticity of supply is defined as a numerical measure of the responsiveness of the quantity supplied of product (A) to a change in price of product (A) alone. It is measured as the percentage change in supply that occurs in response to a percentage change in price.

For example, if, in response to a 10% rise in the price of a good, the quantity supplied increases by 20%, the price elasticity of supply would be $20\%/10\% = 2$.

The quantity of a good supplied can, in the short term, be different from the amount produced, as manufacturers will have stocks which they can build up or run down. In the long run, however, quantity supplied and quantity produced are synonymous.

Promotional Elasticity

Percentage by which the demand will change if the seller's advertising expenses rise by 1%. Most advertising is undertaken by individual sellers to promote their own business. By drawing buyers away from competitors, advertising has a much stronger

effect on the sales of an individual seller than on the market demand. Advertising elasticity of the demand faced by an individual seller tends to be larger than the advertising elasticity of the market demand.

The large increase in promotional elasticity, compared with regular elasticity, is primarily due to (1) brand switching by consumers, (2) inventory behaviour (stockpiling), and (3) transaction utility effects (i.e., the sense of “gain”). By combining household-level data, it was estimated that approximately 80 percent of this increase is attributed to brand switchers.

In the case of a monopoly, brand-switching effects do not exist. In addition, inventory behaviour has an effect on demand only after the consumer has already made at least one purchase during the promotion. Since T_{RP} is roughly equal to the interpurchase time of the product (equation (21)), the effect of inventory behaviour is small in the immediate term and only gains importance in the intermediate term when $t - t_0 = O(T_{RP})$.

Therefore, the increase in demand during the promotional activity of a monopoly is due mostly to transaction utility effects, promotional elasticity is roughly equal to immediate-term elasticity. In a competitive environment, where brand switching does occur, the above relation gives the relative contribution of transaction utility to the increase in promotional elasticity, and it accounts for some of the “missing 20 percent.”

Income Elasticity of Demand

In economics, the income elasticity of demand measures the responsiveness of the quantity demanded of a good to the income of the people demanding the good.

Formula: $(\% \text{change in demand}) / (\% \text{change in income}) = \text{Income elasticity}$

It is measured as the percentage change in demand that occurs in response to a percentage change in income. For example, if, in response to a 10% increase in income, the quantity of a good demanded increased by 20%, the income elasticity of demand would be $20\%/10\% = 2$.

More formally, for a given Marshallian demand function $Q(I, \bar{P})$ for a good is

$$\frac{\partial Q}{\partial I} \frac{I}{Q}$$

With income I , and vector of prices $\bar{P}$.

A negative income elasticity of demand is associated with inferior goods; an increase in income will lead to a fall in the quantity demanded and may lead to changes to more luxurious substitutes.

A positive income elasticity of demand is associated with normal goods; an increase in income will lead to a rise in the quantity demanded. A high positive income elasticity of demand is associated with luxury goods.

A zero income elasticity of demand is an increase in income without leading to a change in the quantity demanded of a good.

Many necessities have an income elasticity of demand between zero and one: expenditure on these goods may increase with income, but not as fast as income does, so the proportion of expenditure on these goods falls as income rises. This observation for food is known as *Engel's law*.

Cross Elasticity of Demand

In economics, the cross elasticity of demand or cross price elasticity of demand measures the responsiveness of the quantity demanded of a good to a change in the price of another good.

It is measured as the percentage change in demand for the first good that occurs in response to a percentage change in price of the second good. For example, if, in response to a 10% increase in the price of fuel, the quantity of new cars that are fuel inefficient demanded decreased by 20%, the cross elasticity of demand would be $-20\%/10\% = -2$.

In the example above, the two goods, fuel and cars, are complements - that is, one is used with the other. In these cases the cross elasticity of demand will be negative. In the case of

perfect complements, the cross elasticity of demand is 'negative' infinity.

Where the two goods are substitutes the cross elasticity of demand will be positive, so that as the price of one goes up the quantity demanded of the other will increase. For example, in response to an increase in the price of fuel, the demand for new cars that are fuel efficient hybrids for example will also rise. In the case of perfect substitutes, the cross elasticity of demand is 'positive' infinity. Where the two goods are independent, the cross elasticity demand will be zero: as the price of one good changes, there will be no change in quantity demanded of the other good. In case of perfect independence, the cross elasticity of demand is zero.

Factors Determining Elasticity of Demand

There are various factors on which elasticity of demand depends:

Nature of the Commodity

In the first place, it depends on the nature of the commodity. Commodities which are supposed to be essential or critical to our daily lives must have an inelastic demand, since price change of these items does not bring about a greater change in quantity demanded.

But, luxury goods have an elastic demand. Demand for these good can be quickly reduced when their prices rise. When their prices fall, consumers demand these goods in larger quantities. However, whether a particular commodity is a necessary or a luxury depends on income, tastes and preferences of the consumer.

A particular good may be necessary to someone having an inelastic demand. Same commodity may be elastic to another consumer. For instance, owning a TV may be a luxury item to a low income person. But the same may be bought as an essential item by a rich person.

Availability of Substitutes

Secondly, commodities having large number of substitutes must have an elastic demand. Some products, such as Horlicks,

Complan, Viva, Maltova, Milo, etc., have quite a large number of close substitutes. A change in the price of, say, Horlicks—the prices of other substitutes remaining constant—will lead a consumer to substitute one beverage for another.

If the price of Horlicks goes down, buyers will demand more of it and less of its substitutes. Conversely, demand is fairly inelastic in the case of those commodities which do not have a large number of substitutes.

Extent of Uses

Thirdly, there are some commodities which can be used for a variety of purposes. For example, electricity. If price per unit of electricity consumed falls, people will reduce their consumption of its substitutes (e.g., coal, gas, etc.) and increase the consumption of electricity.

Coefficient of price elasticity of demand in this case must be greater than one. On the other hand, when a commodity is used only for one or two purposes, a price change will have less effect on its quantity demanded and, therefore, demand will be inelastic.

Habit Good

Fourthly, there are some commodities consumed out of habits and conventions— they have an elastic demand. Even in the face of rising prices of those commodities or falling income, people will consume those (such as, cigarette).

For this reason, price elasticity as well as income elasticity of demand for this type of commodity is inelastic. Further, gold ornaments are used in the marriage ceremony rather out of convention, though gold prices are rising. When gold is used in this way, its demand becomes inelastic.

Time Dimension

Fifthly, shorter the time, lower will be the elasticity of demand. This is because in the short run satisfactory substitutes of a product may not be available. Thus, demand for a product in the short run usually becomes inelastic. Such a commodity will be elastic in the long run when close substitutes may be produced.

Thus, the response of quantity demanded to a change in price will tend to be greater (smaller), the longer (shorter) the time-span considered. In the long run, there is enough time for adjustments to be made following a change in price.

The Importance of being Unimportant

Sixthly, people often pay little attention to the price of a product if it constitutes a relatively small part in their budget. For example, if the price of railway ticket of a tourist who travels by rail once in a year is increased from Rs. 125 to Rs. 135, then he may not postpone his journey. This means he is unresponsive to such price hike and his demand is inelastic. This is called 'the importance of being unimportant'.

Durability

Finally, durable commodities have an elastic demand. If the price of these goods rises, people will spend less on these goods. On the other hand, following a fall in the price of durable commodities (e.g., refrigerator), people demand more of them. In the case of non-durable commodities, demand is elastic.

Importance of the Concept of Elasticity of Demand

The concept of elasticity of demand has both theoretical and practical value. The concept may be used in understanding as well as tackling various economic problems:

Price Determination

Use of the concept of elasticity of demand is required in the price determination of a commodity under different market conditions. Under perfect competition, in the short run in which supply is absolutely inelastic price depends upon the elasticity of demand.

If demand suddenly falls—supply remaining fixed—prices will fall, and, if demand suddenly rises, prices will rise as output cannot be increased. Again, the stability of prices also depends on the elasticity of demand and elasticity of supply. If either the demand or the supply is elastic, fluctuations in prices will be

within narrow limits.

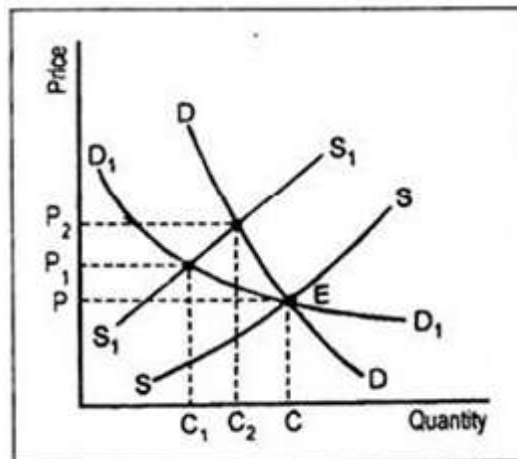
Further, if the demand for an agricultural commodity is inelastic, increased production may spell disaster to the economic condition of farmers. So the government can adopt measures to save the plight of the farmers.

A monopoly seller must have a knowledge relating to the elasticity of demand for his product while determining the price of his commodity.

A monopolist will produce a commodity in the range of his demand curve where demand is said to be elastic. He will never produce in the range of the demand curve where demand is inelastic. Obviously, price determination of the monopoly product will be governed by the elasticity of demand.

Wage Determination

The concept of elasticity of demand is employed in wage determination. Wages, in modern days, are determined through the process of collective bargaining.



Trade union will be successful in raising the wage rate provided labour demand is deemed to be inelastic. This is because of the fact that the degree of substitution between labour and other labour substituting inputs is less. Trade union becomes cautious

in demanding higher wage rates when the demand for labour is said to be elastic.

Under the circumstance, the employer may be forced to employ more machines (assumed to be a cheaper input) than labour. Anyway, this concept may be employed in analysing the problems connected with changes in the conditions of supply. Economists are interested in knowing the effect on employment in the software industry following a rise in the wages of workers engaged in this industry.

Here DD is a rather inelastic demand curve whereas demand curve D_1D_1 is an elastic one.

Both these demand curves intersect the supply curve, SS, at point E. Thus the equilibrium price is OP and equilibrium quantity demanded and supplied is OC.

Let there be an increase in the wages of workers in the computer industry.

Consequently, the supply curve for computer will shift left to S_1S_1 and the price will rise to OP_2 if demand curve is assumed to be DD and to OP_1 if demand curve is D_1D_1 .

However, output contracts more in the case of elastic demand (from OC to OC_1). If demand is inelastic, output will shrink less (from OC to OC_2).

“The general rule is that where demand is elastic, a change in supply will cause the quantity sold to change rather than price; where demand is inelastic, price changes rather than the quantity sold. Thus, trade union will find it more difficult to obtain a wage increase for its members without creating unemployment where the elasticity of demand for the product made is high.” (Jack Harvey)

Policy Determination

The concept of elasticity of demand is of great importance to a finance minister. While imposing tax or raising the existing tax rates, the finance minister must have sufficient knowledge of the elasticity of demand for the taxed commodity. If the demand for the product is inelastic, the purpose of the tax—say revenue—

earning – will be served. That is why taxes are mostly imposed or rates of taxes are raised in the case of commodities having inelastic demand.

Again, the concept may be used in the determination of incidence of a tax. It is easier to shift the burden of taxes on to the consumers if the product demand is assumed to be inelastic. Further, whether exportable or importable be taxed or not, the concept of elasticity may be of great use.

Exchange Rate Determination

In international trade too, the concept may be employed. For instance, as far as exchange rate (i.e., the rate at which one currency is exchanged for another currency) determination is concerned, the concept of elasticity of demand is of great importance.

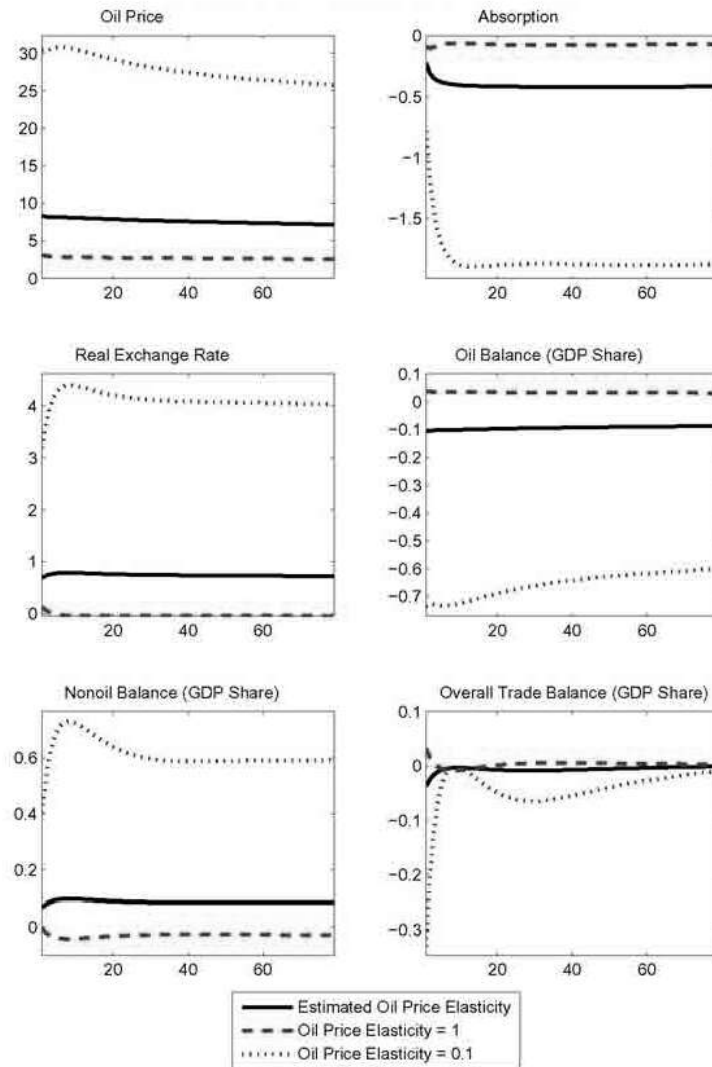
The concept of elasticity of demand is used to justify whether devaluation of a currency is a right step in curbing balance of payment problems of a country. Devaluation is expected to correct the balance of payments disequilibrium if the sum of the elasticities of demand for export and import exceeds unity.

In international trade theory, within the limits set by the comparative costs, the terms of trade also depends on the elasticity of demand of each country for the goods of other countries. In fine, elasticity of demand is a concept which has much applicability as far as business decision-making is concerned and is, therefore, of much importance in modern economics. In fact, most businessmen should try to form as precise an idea of elasticity as possible.

Major Factors which Affects the Elasticity of Demand of a Commodity

Some of the major factors affecting the elasticity of demand of a commodity are as follows:

A change in price does not always lead to the same proportionate change in demand. For example, a small change in price of AC may affect its demand to a considerable extent/ whereas, large change in price of salt may not affect its demand. So, elasticity of demand is different for different goods.



Various factors which affect the elasticity of demand of a commodity are:

Nature of commodity

Elasticity of demand of a commodity is influenced by its

nature. A commodity for a person may be a necessity, a comfort or a luxury.

- i. When a commodity is a necessity like food grains, vegetables, medicines, etc., its demand is generally inelastic as it is required for human survival and its demand does not fluctuate much with change in price.
- ii. When a commodity is a comfort like fan, refrigerator, etc., its demand is generally elastic as consumer can postpone its consumption.
- iii. When a commodity is a luxury like AC, DVD player, etc., its demand is generally more elastic as compared to demand for comforts.
- iv. The term 'luxury' is a relative term as any item (like AC), may be a luxury for a poor person but a necessity for a rich person.

Availability of substitutes

Demand for a commodity with large number of substitutes will be more elastic. The reason is that even a small rise in its prices will induce the buyers to go for its substitutes. For example, a rise in the price of Pepsi encourages buyers to buy Coke and vice-versa.

Thus, availability of close substitutes makes the demand sensitive to change in the prices. On the other hand, commodities with few or no substitutes like wheat and salt have less price elasticity of demand.

Income Level

Elasticity of demand for any commodity is generally less for higher income level groups in comparison to people with low incomes. It happens because rich people are not influenced much by changes in the price of goods. But, poor people are highly affected by increase or decrease in the price of goods. As a result, demand for lower income group is highly elastic.

Level of price

Level of price also affects the price elasticity of demand. Costly

goods like laptop, Plasma TV, etc. have highly elastic demand as their demand is very sensitive to changes in their prices. However, demand for inexpensive goods like needle, match box, etc. is inelastic as change in prices of such goods do not change their demand by a considerable amount.

Postponement of Consumption

Commodities like biscuits, soft drinks, etc. whose demand is not urgent, have highly elastic demand as their consumption can be postponed in case of an increase in their prices. However, commodities with urgent demand like life saving drugs, have inelastic demand because of their immediate requirement.

Number of Uses

If the commodity under consideration has several uses, then its demand will be elastic. When price of such a commodity increases, then it is generally put to only more urgent uses and, as a result, its demand falls. When the prices fall, then it is used for satisfying even less urgent needs and demand rises.

For example, electricity is a multiple-use commodity. Fall in its price will result in substantial increase in its demand, particularly in those uses (like AC, Heat convactor, etc.), where it was not employed formerly due to its high price. On the other hand, a commodity with no or few alternative uses has less elastic demand.

Share in Total Expenditure

Proportion of consumer's income that is spent on a particular commodity also influences the elasticity of demand for it. Greater the proportion of income spent on the commodity, more is the elasticity of demand for it and vice-versa.

Demand for goods like salt, needle, soap, match box, etc. tends to be inelastic as consumers spend a small proportion of their income on such goods.

When prices of such goods change, consumers continue to purchase almost the same quantity of these goods. However, if the proportion of income spent on a commodity is large, then demand for such a commodity will be elastic.

Time Period

Price elasticity of demand is always related to a period of time. It can be a day, a week, a month, a year or a period of several years. Elasticity of demand varies directly with the time period. Demand is generally inelastic in the short period.

It happens because consumers find it difficult to change their habits, in the short period, in order to respond to a change in the price of the given commodity. However, demand is more elastic in long run as it is comparatively easier to shift to other substitutes, if the price of the given commodity rises.

Habits

Commodities, which have become habitual necessities for the consumers, have less elastic demand. It happens because such a commodity becomes a necessity for the consumer and he continues to purchase it even if its price rises. Alcohol, tobacco, cigarettes, etc. are some examples of habit forming commodities.

Finally it can be concluded that elasticity of demand for a commodity is affected by number of factors. However, it is difficult to say, which particular factor or combination of factors determines the elasticity. It all depends upon circumstances of each case.

Factors Influencing Elasticity Of Demand

There are many factors on which the elasticity of demand depends. These factors influence the elasticity of demand of a commodity either individually or cumulatively.

Nature of the Commodity Influence Elasticity of Demand

Why is it that demand for some goods is elastic while the demand for others is inelastic? It mainly depends on the nature of the commodity and the degree of necessity. The elasticity of demand depends on whether a commodity is necessity, comfort or luxury. Normally the demand for necessities of life such as rice, wheat, salt, etc., will be inelastic as these are essential for existence. So, everyone will demand a minimum quantity whatever be the price.

On the other hand the demand for comforts and luxuries may not have inelastic demand. When the prices of these fall, generally, more of the commodities will be demanded. In this discussion we should remember that there is nothing inherent in the quality of a commodity to be called necessity or comfort or luxury. Even in necessities, commodities having substitutes will have elastic demand and commodities having no substitute will have inelastic demand.

Though wheat is a necessity as food for people, a rise in price may make the consumers go in for other cereals. This is not the case with salt which has no substitute. So, the demand for wheat may not be so inelastic as that of demand for salt.

Further in the case of luxuries, it should not be concluded that the elasticity of demand for luxuries will be always large. It depends on the type of luxury. For instance diamonds and articles of jewellery are luxuries used by richer classes. Any minor changes in prices will not affect its demand as these commodities are demanded only by richer people.

So much so, we should make it clear that the elasticity of demand may vary from commodity to commodity and also from group to group. What is luxury to one group may be comfort for another group and necessity for yet another group. Hence the elasticity on the basis of nature of commodity can be studied only on a comparative basis.

Uses of the Commodity influence Elasticity of Demand

If a commodity has only one use, a change in price will not affect the demand much and so it will have inelastic demand. If the commodity has a number of uses, change in price will affect the demand for the commodities in many uses. When a commodity is put to various uses, it will have elastic demand. If the price of that commodity is increased, the commodity will be demanded only in essential uses, and in other uses, substitute materials will be utilized.

For instance a fall in price of coal may make everyone including the householders to demand coal and the demand will be elastic. A rise in price will result in the curtailment of the purchases and

householders will shift to either firewood or oil. Here too, the statement that a commodity having several uses will have elastic demand has to be understood with a restrictive sense. For example coal will have elastic demand in houses but inelastic demand in Railways.

Existence of Substitutes influence Elasticity of Demand

Commodities having substitutes will have elastic demand and goods with no substitutes will have inelastic demand. When the price of a commodity rises, the people would shift their preference to substitute commodities and demand the substitutes with the hope that the price of substitutes will not rise.

Consequently the demand will fall heavily for the commodity for which the price has been increased. Suppose the commodity does not have substitute at all like salt, any change in price will not affect the demand and so the demand will be inelastic.

Postponement of Demand influence Elasticity of Demand

Another important factor affecting the demand in a bigger way is postponement of demand for a commodity. If the demand can be postponed, then the commodity will have elastic demand. If the demand cannot be postponed, it will have inelastic demand.

The demand for rice or medicines cannot be postponed while the demand for mangoes, oranges and apples can be postponed, if the prices of these rise. Hence demand for rice and medicines will be inelastic and the demand for the fruits will be elastic, that is, more will be purchased when the prices come down. Of course this factor, postponement of demand is only a corollary of the kind or nature of commodities already discussed. In the case of necessities, the demand cannot be postponed and so demand becomes inelastic. In the case of commodities which are not necessities, demand can be postponed and so the demand becomes elastic.

Amount of Money Spent influence Elasticity of Demand

Elasticity of demand for a commodity also depends on the proportion of consumer's money spent on the commodity. If the

consumer spends only a little amount on the consumption of a particular commodity, the demand for that will be inelastic.

In the case of items like clothing or food, the consumer spends a large proportion of his income and therefore any increase in price will result in sizeably increasing his total expenditure. So to keep himself fairly within his means, the consumer will reduce the quantity purchased. The demand for these commodities will be elastic.

Habits of Consumers influence Elasticity of Demand

If the consumers are addicted to some habits and customs, then, the demand of the commodity will be inelastic. But if the rise in price persists for a long time, even addicts would try to reduce the demand either by resorting to some alternative substitutes or curbing the habit. Generally, commodities and drugs which are stimulants will have inelastic demand.

Range of Prices of Commodities influence Elasticity of Demand

Elasticity of demand for a commodity depends on the range of prices at which the commodity is sold in the market. At a very high range of prices, the demand will be inelastic; so also at a very low range of prices, the demand will be inelastic.

For example, the price of motor car is in the high range in which only the very rich can buy motor cars. At the price-range of Rs.80,000 any drop in price say by Rs.400 or Rs.500 nor any rise in price by that amount will not affect the demand for cars, since the demand comes from a limited group. At this level, the demand will be inelastic. In the same way, if the price of a commodity is very low, whoever wants to buy, will be able to do so. Any small change in the price at the low level will not affect the demand. At low range prices, the demand will be inelastic, as also at high range of prices. Only in the middle range of prices demands tend to be elastic or moderate.

Time factor in Elasticity influence Elasticity of Demand

Time plays a vital role in the elasticity of demand for a commodity. Demand for commodity exists for a period of time,

say, a day, week, month or year or several years. The supply and demand may also confine to a particular season.

Generally, demand for any commodity will be inelastic during the short period and it will be comparatively elastic during the long period. During the short period, the demand cannot be very responsive to the changes in prices because of the following reasons: —

Suppose the price of a commodity falls, the demand may not immediately rise because it will take sometime for consumers to come to know about the fall in price. Even if they come to know about the fall, the tendency will be to wait and watch for further fall in price. So the response will not be immediate.

Further, during the short time, the consumers may not be able to change their habit or pattern of expenditure. It will take sometime. Moreover there are many goods which are of a durable nature and so even if the prices fall, the demand may not be immediately forthcoming from the existing consumers.

For example, a fall in the price of fountain pens will not make us demand fountain pens unless the pen we are using becomes worn out and useless. So, in the short period whenever there is a fall in price, it may not lead to an increase in the quantity demanded suddenly.

In the same way, a rise in price will not immediately result in the reduction of the quantity demanded during the short time. The consumer will find it very difficult to make adjustments with substitute commodities. A rise in price of coal or firewood used as fuel will not result in the quantity demanded suddenly falling, even though the substitute like kerosene or gas may be cheap.

To make necessary arrangement for a substitute fuel, the consumer should make a capital expenditure like purchasing kerosene stove or gas oven so that the substitute fuel may be used. Initially the demand for coal and firewood will be inelastic. But after a long period, the consumers will make adjustments and have their own stoves, oil or gas, and the demand for coal and firewood will become elastic. So the elasticity of demand is greatly influenced by the time element.

Production Function

MEANING AND TYPES OF PRODUCTION FUNCTION

Why Production should be studied Next. We have made human wants, consumption, and demand the first subjects of our study of economic theory because it is from these that all other economic phenomena take their rise. We have seen why men exert themselves in the work of production.

The next logical step is to inquire how men go about the work of production. We have studied the cause and the laws of demand. We have next to make a similar inquiry regarding supply. Our present study therefore is of the general subject of production," says John Stuart Mill, " of putting things into fit places for being acted upon by their own internal forces, and by those residing in other natural objects, is all that man does or can do with matter."

All Production essentially the Same: It has seemed to some, even among economists of an earlier time, that the farmer is more truly a producer than the manufacturer, and the manufacturer than the merchant; but careful thought discloses the fallacy of such a view. All industrial classes alike produce one or more of the four sorts of utility which we have described, and they do so by changing the relations of things in time or space. The farmer changes the position of grains of corn by dropping them into the earth. Then he removes weeds and throws earth about the rising stalks. Thus man's acts in changing the relations and position of things, aided

by nature's materials and forces, result in more corn for human consumption. The manufacturer in the same way changes the position of pieces of matter, and, aided by natural forces within and without the object of production, he causes matter to assume a form which fits it, or better fits it, for human needs. So, too, the merchant changes the places of things from where they are less useful to where they are more useful, or holds them in one place until a change of external circumstances gives them a greater time utility. He is producing utilities as truly as is the farmer or the manufacturer. Of course it is possible that the utilities actually produced by merchants could be produced with a smaller expenditure of economic force than they are at present, and that saving could be effected by a better organization of the work of production.

Again, it may be that the merchant may now and then secure a larger return for the production of a given quantity of social utility than does the farmer. But all this affords no justification for the popular impression that his work is really less productive in its nature than is that of any other industrial class. The only difference is in the kind of utility that the different classes are engaged in producing. Finally, it must be remembered that in the same way the physician, the teacher, and all others who are engaged in rendering personal services, are creating utilities, and are therefore producers.

Production, then, we may define as the creation of utilities by the application of man's mental and physical powers to the physical universe, which furnishes materials and forces. This application of man's powers we call labour.

We have already defined goods and economic goods: It remains for us here to call attention to the fact that those quantities of utility which result from labour are economic goods, but that not all economic goods are to the same extent the result of labour. One may pick up a diamond or a nugget of gold upon which one has stumbled: in such a case it can hardly be said that the economic good is the result of labour at all. But even in such rare cases it must be remembered that while the one diamond or the one nugget may have required no labour in getting, yet the whole

stock of such goods is the result of toil and suffering and privation for which the value of our diamonds and gold, it is frequently said, does not represent anything like a proper recompense.

There is one clearly marked case of value creation which is not wealth production. The land on which New York and Chicago stand could have been purchased only a few centuries ago for a very small sum of money. The great value which that land now has is to a considerable degree the result of human labour, but much of it is due to the great increase in population, which of itself represents no idea of labour. Such value is a product of social aggregation, not of individual effort. The question of the expediency of allowing individuals to appropriate these individually unearned increments of value will be discussed later. Here it concerns us only to notice that such unearned increments exist; in other words, that there is such a thing in the world as value creation which is not at the same time wealth production.

Individual and Social Wealth: This distinction between the individual and the social standpoint runs all the way through economics, and it is particularly important in the case of the conception of wealth or economic goods. What is wealth to the individual may not be wealth to society, and, on the other hand, what is wealth to society may not be within the ownership of an individual.

Thus a mortgage is wealth to the individual who holds it, but it is not a part of social wealth, since if the claim for which it stands is extinguished, society is neither richer nor poorer. The case is the same with bonds issued by a city, a state, or a nation.

Productive Elements often Overlooked: There are many important facts regarding production which are often overlooked. Thus we are likely to forget that even today a large part of production is household production, and is not designed for the market place at all. The labour of at least half the women of the country is expended in producing material good things for the use of producers.

Again, we are likely to overlook the fact that in the country, where over one-half of the population of the United States lives

and works, there is annually produced a vast amount of goods which are destined not for the market but for home consumption. Vegetables, small fruits, cultivated and wild, butter, eggs, meat, fish caught in public waters, and game are some of the things that occur most readily to the mind. Considerations of this character show the great need of caution in attempting to compare the annual production of one country with that of another, or to compare the annual production of the same country at different periods. Household production is becoming relatively less important, while the production of things for the market, the value of which is readily measured in money, is constantly gaining in importance. Hence, apparent annual production the production of things which have a market price set upon them is increasing more rapidly than is the real annual production. The result is a tendency to overestimate our progress and even to count as progress what may not be progress at all.

Thus, should boarding-house and hotel life displace private housekeeping, annual production might appear to increase as a result of the change, though the real wealth and income of the country would evidently be affected in no such degree. Still further care must be exercised in studying census estimates of wealth. These estimates are ordinarily made in terms of money. Now if commodities are very abundant, their price, other things being equal, will be low, though the real wealth of the country is great. If, for instance, the quantity of cotton cloth produced doubles between two census periods, while the price falls one-half, the total value of the product will appear in the census estimates as equal in the two cases, though it is evident that society in the second period has twice the amount of this valuable commodity.

Defining the Production Function

The production function relates the maximum amount of output that can be obtained from a given number of inputs. In economics, a production function relates physical output of a production process to physical inputs or factors of production. It is a mathematical function that relates the maximum amount of output that can be obtained from a given number of inputs –

generally capital and labor. The production function, therefore, describes a boundary or frontier representing the limit of output obtainable from each feasible combination of inputs.

Firms use the production function to determine how much output they should produce given the price of a good, and what combination of inputs they should use to produce given the price of capital and labor. When firms are deciding how much to produce they typically find that at high levels of production, their marginal costs begin increasing. This is also known as diminishing returns to scale – increasing the quantity of inputs creates a less-than-proportional increase in the quantity of output. If it weren't for diminishing returns to scale, supply could expand without limits without increasing the price of a good.

Increasing marginal costs can be identified using the production function. If a firm has a production function $Q=F(K,L)$ (that is, the quantity of output (Q) is some function of capital (K) and labor (L)), then if $2Q < F(2K,2L)$, the production function has increasing marginal costs and diminishing returns to scale. Similarly, if $2Q > F(2K,2L)$, there are increasing returns to scale, and if $2Q = F(2K,2L)$, there are constant returns to scale.

Examples of Common Production Functions

One very simple example of a production function might be $Q=K+L$, where Q is the quantity of output, K is the amount of capital, and L is the amount of labor used in production. This production function says that a firm can produce one unit of output for every unit of capital or labor it employs. From this production function we can see that this industry has constant returns to scale – that is, the amount of output will increase proportionally to any increase in the amount of inputs.

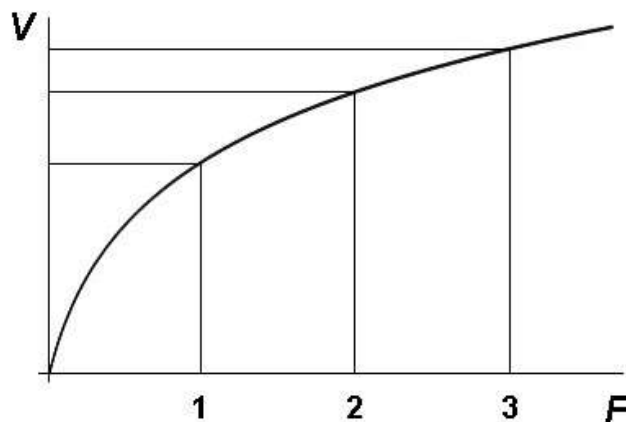
Another common production function is the Cobb-Douglas production function. One example of this type of function is $Q=K^{0.5}L^{0.5}$. This describes a firm that requires the least total number of inputs when the combination of inputs is relatively equal. For example, the firm could produce 25 units of output by using 25 units of capital and 25 of labor, or it could produce the same 25 units of output with 125 units of labor and only one unit of capital.

Finally, the Leontief production function applies to situations in which inputs must be used in fixed proportions; starting from those proportions, if usage of one input is increased without another being increased, output will not change. This production function is given by $Q = \min(K, L)$. For example, a firm with five employees will produce five units of output as long as it has at least five units of capital.

The Law of Diminishing Returns

The law of diminishing returns states that adding more of one factor of production will at some point yield lower per-unit returns.

In economics, diminishing returns (also called diminishing marginal returns) is the decrease in the marginal output of a production process as the amount of a single factor of production is increased, while the amounts of all other factors of production stay constant. The law of diminishing returns states that in all productive processes, adding more of one factor of production, while holding all others constant (“ceteris paribus”), will at some point yield lower per-unit returns. The law of diminishing returns does not imply that adding more of a factor will decrease the total production, a condition known as negative returns, though in fact this is common.



Diminishing Returns: As a factor of production (F) increases, the resulting gain in the volume of output (V) gets smaller and

smaller. For example, the use of fertilizer improves crop production on farms and in gardens; but at some point, adding more and more fertilizer improves the yield less per unit of fertilizer, and excessive quantities can even reduce the yield. A common sort of example is adding more workers to a job, such as assembling a car on a factory floor. At some point, adding more workers causes problems such as workers getting in each other's way or frequently finding themselves waiting for access to a part. In all of these processes, producing one more unit of output will eventually cost increasingly more, due to inputs being used less and less effectively.

This increase in the marginal cost of output as production increases can be graphed as the marginal cost curve, with quantity of output on the x axis and marginal cost on the y axis. For many firms, the marginal cost curve will initially be downward sloping, representing added efficiency as production increases. If the law of diminishing returns holds, however, the marginal cost curve will eventually slope upward and continue to rise, representing the higher and higher marginal costs associated with additional output.

The Law of Diminishing Returns and Average Cost

The average total cost of production is the total cost of producing all output divided by the number of units produced. For example, if the car factory can produce 20 cars at a total cost of \$200,000, the average cost of production is \$10,000. Average total cost is interpreted as the cost of a typical unit of production. So in our example each of the 20 cars produced had a typical cost per unit of \$10,000. Average total cost can also be graphed with quantity of output on the x axis and average cost on the y-axis.

What will this average total cost curve look like? In the short run, a firm has a set amount of capital and can only increase or decrease production by hiring more or less labor. The fixed costs of capital are high, but the variable costs of labor are low, so costs increase more slowly than output as production increases. As long as the marginal cost of production is lower than the average total cost of production, the average cost is decreasing. However, as marginal costs increase due to the law of diminishing returns, the marginal cost of production will eventually be higher than the

average total cost and the average cost will begin to increase. The short run average total cost curve (SRAC) will therefore be U-shaped for most firms.

The long-run average cost curve (LRAC) depicts the cost per unit of output in the long run – that is, when all productive inputs' usage levels can be varied. The typical LRAC curve is also U-shaped but for different reasons: it reflects increasing returns to scale where negatively-sloped, constant returns to scale where horizontal, and decreasing returns (due to increases in factor prices) where positively sloped.

Concept of Production Functions

In micro-economics, a production function is a function that specifies the output of a firm for all combinations of inputs. A meta-production function (sometimes metaproduction function) compares the practice of the existing entities converting inputs into output to determine the most efficient practice production function of the existing entities, whether the most efficient feasible practice production or the most efficient actual practice production. In either case, the maximum output of a technologically-determined production process is a mathematical function of one or more inputs. Put another way, given the set of all technically feasible combinations of output and inputs, only the combinations encompassing a maximum output for a specified set of inputs would constitute the production function. Alternatively, a production function can be defined as the specification of the minimum input requirements needed to produce designated quantities of output, given available technology. It is usually presumed that unique production functions can be constructed for every production technology.

By assuming that the maximum output technologically possible from a given set of inputs is achieved, economists using a production function in analysis are abstracting from the engineering and managerial problems inherently associated with a particular production process.

The engineering and managerial problems of technical efficiency are assumed to be solved, so that analysis can focus on

the problems of allocative efficiency. The firm is assumed to be making allocative choices concerning how much of each input factor to use and how much output to produce, given the cost (purchase price) of each factor, the selling price of the output, and the technological determinants represented by the production function. A decision frame in which one or more inputs are held constant may be used; for example, (physical) capital may be assumed to be fixed (constant) in the short run, and labour and possibly other inputs such as raw materials variable, while in the long run, the quantities of both capital and the other factors that may be chosen by the firm are variable. In the long run, the firm may even have a choice of technologies, represented by various possible production functions.

The relationship of output to inputs is non-monetary; that is, a production function relates physical inputs to physical outputs, and prices and costs are not reflected in the function. But the production function is not a full model of the production process: it deliberately abstracts from inherent aspects of physical production processes that some would argue are essential, including error, entropy or waste. Moreover, production functions do not ordinarily model the business processes, either, ignoring the role of management.

The primary purpose of the production function is to address allocative efficiency in the use of factor inputs in production and the resulting distribution of income to those factors. Under certain assumptions, the production function can be used to derive a marginal product for each factor, which implies an ideal division of the income generated from output into an income due to each input factor of production.

Specifying the Production Function

A production function can be expressed in a functional form as the right side of

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

where:

Q = quantity of output

$X_1, X_2, X_3, \dots, X_n$ = quantities of factor inputs (such as capital, labour, land or raw materials).

If Q is not a matrix (i.e. a scalar, a vector, or even a diagonal matrix), then this form does not encompass joint production, which is a production process that has multiple co-products. On the other hand, if f maps from R^n to R^k then it is a joint production function expressing the determination of k different types of output based on the joint usage of the specified quantities of the n inputs.

One formulation, unlikely to be relevant in practice, is as a linear function:

$$Q = a + bX_1 + cX_2 + dX_3 + \dots$$

where a, b, c , and d are parameters that are determined empirically.

Another is as a Cobb-Douglas production function:

$$Q = aX_1^b X_2^c \dots$$

The Leontief production function applies to situations in which inputs must be used in fixed proportions; starting from those proportions, if usage of one input is increased without another being increased, output will not change. This production function is given by

$$Q = \min(aX_1, bX_2, \dots).$$

Other forms include the constant elasticity of substitution production function (CES), which is a generalized form of the Cobb-Douglas function, and the quadratic production function. The best form of the equation to use and the values of the parameters ($a, b, c, \dots$) vary from company to company and industry to industry. In a short run production function at least one of the X 's (inputs) is fixed. In the long run all factor inputs are variable at the discretion of management.

Any of these equations can be plotted on a graph. A typical (quadratic) production function is shown in the following diagram under the assumption of a single variable input (or fixed ratios of inputs so the can be treated as a single variable). All points above the production function are unobtainable with current technology, all points below are technically feasible, and all points

on the function show the maximum quantity of output obtainable at the specified level of usage of the input. From the origin, through points A, B, and C, the production function is rising, indicating that as additional units of inputs are used, the quantity of output also increases. Beyond point C, the employment of additional units of inputs produces no additional output (in fact, total output starts to decline); the variable input is being used too intensively. With too much variable input use relative to the available fixed inputs, the company is experiencing negative marginal returns to variable inputs, and diminishing total returns. In the diagram this is illustrated by the negative marginal physical product curve (MPP) beyond point Z, and the declining production function beyond point C.

From the origin to point A, the firm is experiencing increasing returns to variable inputs: As additional inputs are employed, output increases at an increasing rate. Both marginal physical product (MPP, the derivative of the production function) and average physical product (APP, the ratio of output to the variable input) are rising. The inflection point A defines the point beyond which there are diminishing marginal returns, as can be seen from the declining MPP curve beyond point X. From point A to point C, the firm is experiencing positive but decreasing marginal returns to the variable input. As additional units of the input are employed, output increases but at a decreasing rate. Point B is the point beyond which there are diminishing average returns, as shown by the declining slope of the average physical product curve (APP) beyond point Y. Point B is just tangent to the steepest ray from the origin hence the average physical product is at a maximum. Beyond point B, mathematical necessity requires that the marginal curve must be below the average curve.

Stages of Production

To simplify the interpretation of a production function, it is common to divide its range into 3 stages. In Stage 1 (from the origin to point B) the variable input is being used with increasing output per unit, the latter reaching a maximum at point B (since the average physical product is at its maximum at that point).

Because the output per unit of the variable input is improving throughout stage 1, a price-taking firm will always operate beyond this stage. In Stage 2, output increases at a decreasing rate, and the average and marginal physical product are declining.

However, the average product of fixed inputs (not shown) is still rising, because output is rising while fixed input usage is constant. In this stage, the employment of additional variable inputs increases the output per unit of fixed input but decreases the output per unit of the variable input. The optimum input/output combination for the price-taking firm will be in stage 2, although a firm facing a downward-sloped demand curve might find it most profitable to operate in Stage 1. In Stage 3, too much variable input is being used relative to the available fixed inputs: variable inputs are over-utilized in the sense that their presence on the margin obstructs the production process rather than enhancing it. The output per unit of both the fixed and the variable input declines throughout this stage. At the boundary between stage 2 and stage 3, the highest possible output is being obtained from the fixed input.

Shifting a Production Function

By definition, in the long run the firm can change its scale of operations by adjusting the level of inputs that are fixed in the short run, thereby shifting the production function upward as plotted against the variable input. If fixed inputs are lumpy, adjustments to the scale of operations may be more significant than what is required to merely balance production capacity with demand. For example, you may only need to increase production by a million units per year to keep up with demand, but the production equipment upgrades that are available may involve increasing productive capacity by 2 million units per year.

If a firm is operating at a profit-maximizing level in stage one, it might, in the long run, choose to reduce its scale of operations (by selling capital equipment). By reducing the amount of fixed capital inputs, the production function will shift down. The beginning of stage 2 shifts from B1 to B2. The (unchanged) profit-maximizing output level will now be in stage 2.

Homogeneous and Homothetic Production Functions

There are two special classes of production functions that are often analysed. The production function $Q = f(X_1, X_2)$ is said to be homogeneous of degree n , if given any positive constant k , $f(kX_1, kX_2) = k^n f(X_1, X_2)$. If $n > 1$, the function exhibits increasing returns to scale, and it exhibits decreasing returns to scale if $n < 1$. If it is homogeneous of degree 1, it exhibits constant returns to scale.

The presence of increasing returns means that a one percent increase in the usage levels of all inputs would result in a greater than one percent increase in output; the presence of decreasing returns means that it would result in a less than one percent increase in output. Constant returns to scale is the in-between case. In the Cobb-Douglas production function referred to above, returns to scale are increasing if $b + c + \dots > 1$, decreasing if $b + c + \dots < 1$, and constant if $b + c + \dots = 1$.

If a production function is homogeneous of degree one, it is sometimes called "linearly homogeneous". A linearly homogeneous production function with inputs capital and labour has the properties that the marginal and average physical products of both capital and labour can be expressed as functions of the capital-labour ratio alone. Moreover, in this case if each input is paid at a rate equal to its marginal product, the firm's revenues will be exactly exhausted and there will be no excess economic profit. Homothetic functions are functions whose marginal technical rate of substitution (the slope of the isoquant, a curve drawn through the set of points in say labour-capital space at which the same quantity of output is produced for varying combinations of the inputs) is homogeneous of degree zero. Due to this, along rays coming from the origin, the slopes of the isoquants will be the same. Homothetic functions are of the form $F(h(X_1, X_2))$ where $F(y)$ is a monotonically increasing function (the derivative of $F(y)$ is positive ($dF/dy > 0$)), and the function $h(X_1, X_2)$ is a homogeneous function of any degree.

Aggregate Production Functions

In macroeconomics, aggregate production functions for whole nations are sometimes constructed. In theory they are the

summation of all the production functions of individual producers; however there are methodological problems associated with aggregate production functions, and economists have debated extensively whether the concept is valid.

Criticisms of Production Functions

There are two major criticisms of the standard form of the production function.

On the Concept of Capital

During the 1950s, '60s, and '70s there was a lively debate about the theoretical soundness of production functions. Although the criticism was directed primarily at aggregate production functions, microeconomic production functions were also put under scrutiny.

The debate began in 1953 when Joan Robinson criticized the way the factor input capital was measured and how the notion of factor proportions had distracted economists.

According to the argument, it is impossible to conceive of capital in such a way that its quantity is independent of the rates of interest and wages.

The problem is that this independence is a precondition of constructing an isoquant. Further, the slope of the isoquant helps determine relative factor prices, but the curve cannot be constructed (and its slope measured) unless the prices are known beforehand.

On the Empirical Relevance

As a result of the criticism on their weak theoretical grounds, it has been claimed that empirical results firmly support the use of neoclassical *well behaved* aggregate production functions. Nevertheless, Anwar Shaikh has demonstrated that they also has no empirical relevance, as long as alleged good fit outcomes from an accounting identity, not from any underlying laws of production/distribution.

Natural Resources

Often natural resources are omitted from production functions. When Solow and Stiglitz sought to make the production function

more realistic by adding in natural resources, they did it in a manner that economist Georgescu-Roegen criticized as a “conjuring trick” that failed to address the laws of thermodynamics, since their variant allows capital and labour to be infinitely substituted for natural resources. Neither Solow nor Stiglitz addressed his criticism, despite an invitation to do so in the September 1997 issue of the journal *Ecological Economics*.

LAWS OF VARIABLE PROPORTIONS

Variable Proportions

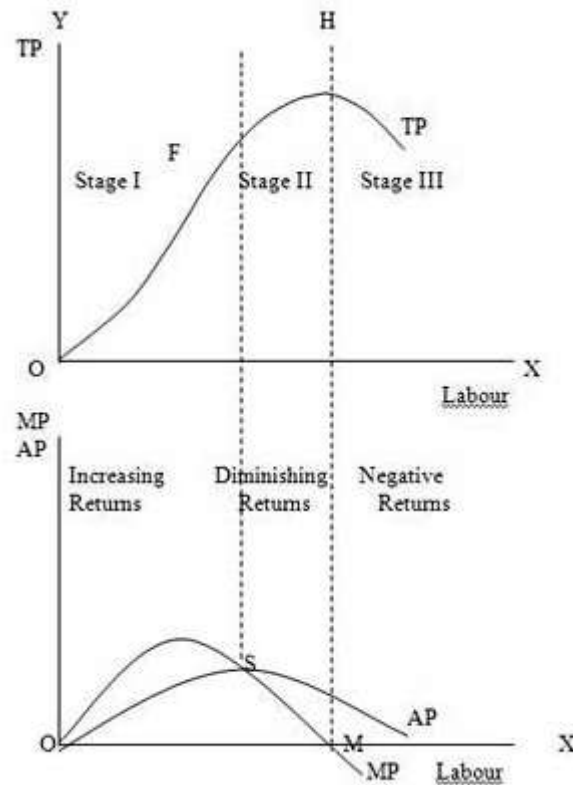
The law of variable proportions states that as the quantity of one factor is increased, keeping the other factors fixed, the marginal product of that factor will eventually decline. This means that upto the use of a certain amount of variable factor, marginal product of the factor may increase and after a certain stage it starts diminishing. When the variable factor becomes relatively abundant, the marginal product may become negative.

Assumptions: The law of variable proportions holds good under the following conditions:

1. *Constant State of Technology*: First, the state of technology is assumed to be given and unchanged. If there is improvement in the technology, then the marginal product may rise instead of diminishing.
2. *Fixed Amount of Other Factors*: Secondly, there must be some inputs whose quantity is kept fixed. It is only in this way that we can alter the factor proportions and know its effects on output. The law does not apply if all factors are proportionately varied.
3. *Possibility of Varying the Factor proportions*: Thirdly, the law is based upon the possibility of varying the proportions in which the various factors can be combined to produce a product. The law does not apply if the factors must be used in fixed proportions to yield a product.

Three Stages of the Law of Variable Proportions: These stages are illustrated is measured on the X-axis and output on the Y-axis.

Stage 1. Stage of Increasing Returns: In this stage, total product increases at an increasing rate up to a point.



This is because the efficiency of the fixed factors increases as additional units of the variable factors are added to it. From the origin to the point F, slope of the total product curve TP is increasing i.e. the curve TP is concave upwards upto the point F, which means that the marginal product MP of labour rises. The point F where the total product stops increasing at an increasing rate and starts increasing at a diminishing rate is called the point of inflection. Corresponding vertically to this point of inflection marginal product of labour is maximum, after which it diminishes. This stage is called the stage of increasing returns because the average product of the variable factor increases throughout this stage. This stage

ends at the point where the average product curve reaches its highest point.

Stage 2. Stage of Diminishing Returns: In this stage, total product continues to increase but at a diminishing rate until it reaches its maximum point H where the second stage ends. In this stage both the marginal product and average product of labour are diminishing but are positive. This is because the fixed factor becomes inadequate relative to the quantity of the variable factor. At the end of the second stage, i.e., at point M marginal product of labour is zero which corresponds to the maximum point H of the total product curve TP. This stage is important because the firm will seek to produce in this range.

Stage 3. Stage of Negative Returns: In stage 3, total product declines and therefore the TP curve slopes downward. As a result, marginal product of labour is negative and the MP curve falls below the X-axis. In this stage the variable factor (labour) is too much relative to the fixed factor.

Importance and Applicability of the Law of Variable Proportion

The Law of Variable Proportion has universal applicability in any branch of production. It forms the basis of a number of doctrines in economics. The Malthusian theory of population stems from the fact that food supply does not increase faster than the growth in population because of the operation of the law of diminishing returns in agriculture.

Ricardo also based his theory of rent on this principle. According to him rent arises because the operation of the law of diminishing return forces the application of additional doses of labour and capital on a piece of land. Similarly the law of diminishing marginal utility and that of diminishing marginal physical productivity in the theory of distribution are also based on this theory.

The law is of fundamental importance for understanding the problems of underdeveloped countries. In such agricultural economies the pressure of population on land increases with the

increase in population. This leads to declining or even zero or negative marginal productivity of workers. This explains the operation of the law of diminishing returns in LDCs in its intensive form. Ragnar Nurkse have suggested ways to make use of these disguisedly unemployed labour by withdrawing them and putting them in those occupations where the marginal productivity is positive.

Definitions

“As the proportion of the factor in a combination of factors is increased after a point, first the marginal and then the average product of that factor will diminish.” Benham

“An increase in some inputs relative to other fixed inputs will in a given state of technology cause output to increase, but after a point the extra output resulting from the same additions of extra inputs will become less and less.” Samuelson

“The law of variable proportion states that if the inputs of one resource is increased by equal increment per unit of time while the inputs of other resources are held constant, total output will increase, but beyond some point the resulting output increases will become smaller and smaller.” Leftwitch

Assumptions

Law of variable proportions is based on following assumptions:

- (i) Constant Technology: The state of technology is assumed to be given and constant. If there is an improvement in technology the production function will move upward.
- (ii) Factor Proportions are Variable: The law assumes that factor proportions are variable. If factors of production are to be combined in a fixed proportion, the law has no validity.
- (iii) Homogeneous Factor Units: The units of variable factor are homogeneous. Each unit is identical in quality and amount with every other unit.
- (iv) Short-Run: The law operates in the short-run when it is not possible to vary all factor inputs.

Explanation of the Law

In order to understand the law of variable proportions we take the example of agriculture. Suppose land and labour are the only two factors of production.

By keeping land as a fixed factor, the production of variable factor i.e., labour can be shown with the help of the following table:

Table 1.

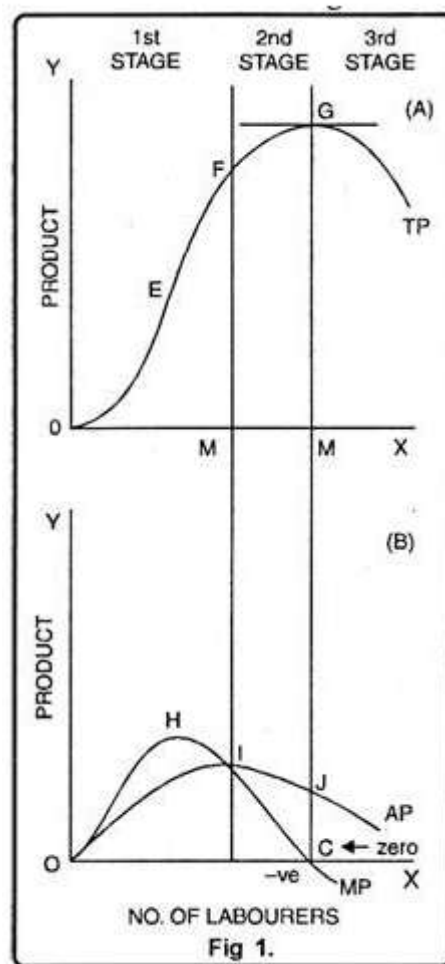
Units of Land	Units of Labour	Total Production	Average Production	Marginal Production
10 Acres	0	—	—	—
"	1	20	20	20
"	2	50	25	30
"	3	90	30	40
"	4	120	30	30
"	5	140	28	20
"	6	150	25	10
"	7	150	21.3	0
"	8	140	17.5	-10

There are three stages of the law of variable proportion. In the first stage average production increases as there are more and more doses of labour and capital employed with fixed factors (land). We see that total product, average product, and marginal product increases but average product and marginal product increases up to 40 units. Later on, both start decreasing because proportion of workers to land was sufficient and land is not properly used. This is the end of the first stage.

The second stage starts from where the first stage ends or where $AP=MP$. In this stage, average product and marginal product start falling. We should note that marginal product falls at a faster rate than the average product. Here, total product increases at a diminishing rate. It is also maximum at 70 units of labour where marginal product becomes zero while average product is never zero or negative. The third stage begins where second stage ends. This starts from 8th unit. Here, marginal product is negative and total product falls but average product is still positive. At this stage, any additional dose leads to positive nuisance because additional dose leads to negative marginal product.

Graphic Presentation

OX axis, we have measured number of labourers while quantity of product is shown on OY axis.



TP is total product curve. Up to point 'E', total product is increasing at increasing rate. Between points E and G it is increasing at the decreasing rate. Here marginal product has started falling. At point 'G' i.e., when 7 units of labourers are employed, total product is maximum while, marginal product is zero. Thereafter,

it begins to diminish corresponding to negative marginal product. In the lower part of the figure MP is marginal product curve.

Up to point 'H' marginal product increases. At point 'H', i.e., when 3 units of labourers are employed, it is maximum. After that, marginal product begins to decrease. Before point 'T' marginal product becomes zero at point C and it turns negative. AP curve represents average product. Before point 'T', average product is less than marginal product. At point 'T' average product is maximum. Up to point T, average product increases but after that it starts to diminish.

Three Stages of the Law

First Stage

First stage starts from point 'O' and ends up to point F. At point F average product is maximum and is equal to marginal product. In this stage, total product increases initially at increasing rate up to point E. between 'E' and 'F' it increases at diminishing rate. Similarly marginal product also increases initially and reaches its maximum at point 'H'. Later on, it begins to diminish and becomes equal to average product at point T. In this stage, marginal product exceeds average product ($MP > AP$).

Second Stage

It begins from the point F. In this stage, total product increases at diminishing rate and is at its maximum at point 'G' correspondingly marginal product diminishes rapidly and becomes 'zero' at point 'C'. Average product is maximum at point 'T' and thereafter it begins to decrease. In this stage, marginal product is less than average product ($MP < AP$).

Third Stage

This stage begins beyond point 'G'. Here total product starts diminishing. Average product also declines. Marginal product turns negative. Law of diminishing returns firmly manifests itself. In this stage, no firm will produce anything. This happens because marginal product of the labour becomes negative. The employer will suffer losses by employing more units of labourers. However,

of the three stages, a firm will like to produce up to any given point in the second stage only.

Total Product	Marginal Product	Average Product
Stage I First increases at increasing rate then at diminishing rate.	Increases in the beginning then reaches a maximum and begins to decrease.	First increases, continues to increase and becomes maximum.
Stage II Continues to increase at diminishing rate and becomes maximum.	Continues to diminish and becomes equal to zero.	Becomes equal to MP and then begins to diminish.
Stage III Diminishes	Becomes negative.	Continues to diminish but will always be greater than zero.

In Which Stage Rational Decision is Possible

To make the things simple, let us suppose that, a is variable factor and b is the fixed factor. And $a_1, a_2, a_3, \dots$ are units of a and $b_1, b_2, b_3, \dots$ are unit of b .

Stage I is characterized by increasing AP, so that the total product must also be increasing. This means that the efficiency of the variable factor of production is increasing i.e., output per unit of a is increasing. The efficiency of b , the fixed factor, is also increasing, since the total product with b_1 is increasing.

The stage II is characterized by decreasing AP and a decreasing MP, but with MP not negative. Thus, the efficiency of the variable factor is falling, while the efficiency of b , the fixed factor, is increasing, since the TP with b_1 continues to increase.

Finally, stage III is characterized by falling AP and MP, and further by negative MP. Thus, the efficiency of both the fixed and variable factor is decreasing.

Rational Decision

Stage II becomes the relevant and important stage of production. Production will not take place in either of the other two stages. It means production will not take place in stage III and stage I. Thus, a rational producer will operate in stage II.

Suppose *b* were a free resource; i.e., it commanded no price. An entrepreneur would want to achieve the greatest efficiency possible from the factor for which he is paying, i.e., from factor *a*.

Thus, he would want to produce where AP is maximum or at the boundary between stage I and II.

If on the other hand, *a* were the free resource, then he would want to employ *b* to its most efficient point; this is the boundary between stage II and III.

Obviously, if both resources commanded a price, he would produce somewhere in stage II. At what place in this stage production takes place would depend upon the relative prices of *a* and *b*.

Condition or Causes of Applicability

There are many causes which are responsible for the application of the law of variable proportions.

They are as follows:

1. Under Utilization of Fixed Factor: In initial stage of production, fixed factors of production like land or machine, is under-utilized. More units of variable factor, like labour, are needed for its proper utilization. As a result of employment of additional units of variable factors there is proper utilization of fixed factor. In short, increasing returns to a factor begins to manifest itself in the first stage.
2. Fixed Factors of Production. The foremost cause of the operation of this law is that some of the factors of production are fixed during the short period. When the fixed factor is used with variable factor, then its ratio compared to variable factor falls. Production is the result of the co-operation of all factors. When an additional unit of a variable factor has to produce with the help of relatively fixed factor, then the marginal return of variable factor begins to decline.
3. Optimum Production: After making the optimum use of a fixed factor, then the marginal return of such variable

factor begins to diminish. The simple reason is that after the optimum use, the ratio of fixed and variable factors become defective. Let us suppose a machine is a fixed factor of production. It is put to optimum use when 4 labourers are employed on it. If 5 labourers are put on it, then total production increases very little and the marginal product diminishes.

4. Imperfect Substitutes: Mrs. Joan Robinson has put the argument that imperfect substitution of factors is mainly responsible for the operation of the law of diminishing returns. One factor cannot be used in place of the other factor. After optimum use of fixed factors, variable factors are increased and the amount of fixed factor could be increased by its substitutes.

Such a substitution would increase the production in the same proportion as earlier. But in real practice factors are imperfect substitutes. However, after the optimum use of a fixed factor, it cannot be substituted by another factor.

Applicability of the Law of Variable Proportions

The law of variable proportions is universal as it applies to all fields of production. This law applies to any field of production where some factors are fixed and others are variable. That is why it is called the law of universal application.

The main cause of application of this law is the fixity of any one factor. Land, mines, fisheries, and house building etc. are not the only examples of fixed factors.

Machines, raw materials may also become fixed in the short period. Therefore, this law holds good in all activities of production etc. agriculture, mining, manufacturing industries.

1. Application to Agriculture: With a view of raising agricultural production, labour and capital can be increased to any extent but not the land, being fixed factor. Thus when more and more units of variable factors like labour and capital are applied to a fixed factor then their marginal product starts to diminish and this law becomes operative.

2. Application to Industries: In order to increase production of manufactured goods, factors of production has to be increased. It can be increased as desired for a long period, being variable factors. Thus, law of increasing returns operates in industries for a long period. But, this situation arises when additional units of labour, capital and enterprise are of inferior quality or are available at higher cost.

As a result, after a point, marginal product increases less proportionately than increase in the units of labour and capital. In this way, the law is equally valid in industries.

Postponement of the Law

The postponement of the law of variable proportions is possible under following conditions:

- (i) Improvement in Technique of Production: The operation of the law can be postponed in case variable factors techniques of production are improved.
- (ii) Perfect Substitute: The law of variable proportion can also be postponed in case factors of production are made perfect substitutes i.e., when one factor can be substituted for the other.

Law of Variable Proportions: Meaning, Definition, Assumption and Stages

Law of variable proportions occupies an important place in economic theory. This law examines the production function with one factor variable, keeping the quantities of other factors fixed. In other words, it refers to the input-output relation when output is increased by varying the quantity of one input.

When the quantity of one factor is varied, keeping the quantity of other factors constant, the proportion between the variable factor and the fixed factor is altered; the ratio of employment of the variable factor to that of the fixed factor goes on increasing as the quantity of the variable factor is increased.

Since under this law we study the effects on output of variation in factor proportions, this is also known as the law of variable

proportions. Thus law of variable proportions is the new name for the famous “Law of Diminishing Returns” of classical economics. This law has played a vital role in the history of economic thought and occupies an equally important place in modern economic theory. This law has been supported by the empirical evidence about the real world. The law of variable proportions or diminishing returns has been stated by various economists in the following manner: As equal increments of one input are added; the inputs of other productive services being held constant, beyond a certain point the resulting increments of product will decrease, i.e., the marginal products will diminish,” (G. Stigler)

“As the proportion of one factor in a combination of factors is increased, after a point, first the marginal and then the average product of that factor will diminish.” (F. Benham)

“An increase in some inputs relative to other fixed inputs will, in a given state of technology, cause output to increase; but after a point the extra output resulting from the same addition of extra inputs will become less.” (Paul A. Samuelson)

Marshall discussed the law of diminishing returns in relation to agriculture. He defines the law as follows: “An increase in the capital and labour applied in the cultivation of land causes in general a less than proportionate increase in the amount of product raised unless it happens to coincide with an improvement in the arts of agriculture.”

It is obvious from the above definitions of the law of variable proportions (or the law of diminishing returns) that it refers to the behaviour of output as the quantity of one factor is increased, keeping the quantity of other factors fixed and further it states that the marginal product and average product will eventually decline.

Assumptions of the Law

The law of variable proportions or diminishing returns, as stated above, holds good under the following conditions:

1. First, the state of technology is assumed to be given and unchanged. If there is improvement in the technology, then marginal and average products may rise instead of diminishing.

2. Secondly, there must be some inputs whose quantity is kept fixed. This is one of the ways by which we can alter the factor proportions and know its effect on output. This law does not apply in case all factors are proportionately varied. Behaviour of output as a result of the variation in all inputs is discussed under "returns to scale".
3. Thirdly the law is based upon the possibility of varying the proportions in which the various factors can be combined to produce a product. The law does not apply to those cases where the factors must be used in fixed proportions to yield a product.

When the various factors are required to be used in rigidly fixed proportions, then the increase in one factor would not lead to any increase in output, that is, the marginal product of the factor will then be zero and not diminishing. It may, however, be pointed out that products requiring fixed proportions of factors are quite uncommon. Thus, the law of variable proportion applies to most of the cases of production in the real world.

<i>Units of Labour</i>	<i>Total Product (Quintals)</i>	<i>Marginal Product (Quintals)</i>	<i>Average Product (Quintals)</i>
L	Q	$\frac{\Delta Q}{\Delta L}$	$\frac{Q}{L}$
1	80	80	80
2	170	90	85
3	270	100	90
4	368	98	92
5	430	62	86
6	480	50	80
7	504	24	72
8	504	0	63
9	495	-9	55
10	480	-15	48

The law of variable proportions. We shall first explain it by considering. Assume that there is a given fixed amount of land, with which more units of the variable factor labour, is used to produce agricultural output.

With a given fixed quantity of land, as a farmer raises employment of labour from one unit to 7 units, the total product

increases from 80 quintals to 504 quintals of wheat. Beyond the employment of 8 units of labour, total product diminishes. It is worth noting that up to the use of 3 units of labour, total product increases at an increasing rate.

This fact is clearly revealed from column 3 which shows successive marginal products of labour as extra units of labour are used. Marginal product of labour, it may be recalled, is the increment in total output due to the use of an extra unit of labour.

The marginal product of labour initially rises and beyond the use of three units of labour, it starts diminishing. Thus when 3 units of labour are employed, marginal product of labour is 100 and with the use of 4th and 5th units of labour marginal product of labour falls to 98 and 62 respectively.

Beyond the use of eight units of labour, total product diminishes and therefore marginal product of labour becomes negative. As regards average product of labour, it rises up to the use of fourth unit of labour and beyond that it is falling through-out.

Three Stages of the Law of Variable Proportions

The behaviour of output when the varying quantity of one factor is combined with a fixed quantity of the other can be divided into three distinct stages. In order to understand these three stages it is better to graphically illustrate the production function with one factor variable. The X-axis the quantity of the variable factor is measured and on the F-axis the total product, average product and marginal product are measured. How the total product, average product and marginal product of a variable factor change as a result of the increase in its quantity, that is, by increasing the quantity of one factor to a fixed quantity of the others.

The total product curve TP of variable factor goes on increasing to a point and after that it starts declining. In the bottom panel average and marginal product curves of labour also rise and then decline; marginal product curve starts declining earlier than the average product curve.

The behaviour of these total, average and marginal products of the variable factor as a result of the increase in its amount is generally divided into three stages which are explained below:

Stage 1: In this stage, total product curve TP increases at an increasing rate up to a point. From the origin to the point F, slope of the total product curve TP is increasing, that is, up to the point F, the total product increases at an increasing rate (the total product curve TP is concave upward upto the point F), which means that the marginal product MP of the variable factor is rising.

From the point F onwards during the stage 1, the total product curve goes on rising but its slope is declining which means that from point F onwards the total product increases at a diminishing rate (total product curve TP is concave down-ward), i.e., marginal product falls but is positive.

The point F where the total product stops increasing at an increasing rate and starts increasing at the diminishing rate is called the point of inflection. Vertically corresponding to this point of inflection marginal product is maximum, after which it starts diminishing.

Thus, marginal product of the variable factor starts diminishing beyond OL amount of the variable factor. That is, law of diminishing returns starts operating in stage 1 from point D on the MP curve or from OL amount of the variable factor used.

This first stage ends where the average product curve AP reaches its highest point, that is, point S on AP curve or CW amount of the variable factor used. During stage 1, when marginal product of the variable factor is falling it still exceeds its average product and so continues to cause the average product curve to rise.

Thus, during stage 1, whereas marginal product curve of a variable factor rises in a part and then falls, the average product curve rises throughout. In the first stage, the quantity of the fixed factor is too much relative to the quantity of the variable factor so that if some of the fixed factor is withdrawn, the total product will increase. Thus, in the first stage marginal product of the fixed factor is negative.

Stage 2: In stage 2, the total product continues to increase at a diminishing rate until it reaches its maximum point H where the second stage ends. In this stage both the marginal product and the

average product of the variable factor are diminishing but remain positive.

At the end of the second stage, that is, at point M marginal product of the variable factor is zero (corresponding to the highest point H of the total product curve TP). Stage 2 is very crucial and important because as will be explained below the firm will seek to produce in its range.

Stage 3: Stage of Negative Returns: In stage 3 with the increase in the variable factor the total product declines and therefore the total product curve TP slopes downward. As a result, marginal product of the variable factor is negative and the marginal product curve MP goes below the X-axis. In this stage the variable factor is too much relative to the fixed factor. This stage is called the stage of negative returns, since the marginal product of the variable factor is negative during this stage.

It may be noted that stage 1 and stage 3 are completely symmetrical. In stage 1 the fixed factor is too much relative to the variable factor. Therefore, in stage 1, marginal product of the fixed factor is negative. On the other hand, in stage 3 the variable factor is too much relative to the fixed factor. Therefore, in stage 3, the marginal product of the variable factor is negative.

The Stage of Operation

Now, an important question is in which stage a rational producer will seek to produce. A rational producer will never choose to produce in stage 3 where marginal product of the variable factor is negative. Marginal product of the variable factor being negative in stage 3, a producer can always increase his output by reducing the amount of the variable factor.

It is thus clear that a rational producer will never be producing in stage 3. Even if the variable factor is free, the rational producer will stop at the end of the second stage where the marginal product of the variable factor is zero.

At the end point M of the second stage where the marginal product of the variable factor is zero, the producer will be maximising the total product and will thus be making maximum use of the variable factor. A rational producer will also not choose

to produce in stage 1 where the marginal product of the fixed factor is negative.

A producer producing in stage 1 means that he will not be making the best use of the fixed factor and further that he will not be utilising fully the opportunities of increasing production by increasing quantity of the variable factor whose average product continues to rise throughout the stage 1. Thus, a rational entrepreneur will not stop in stage 1 but will expand further.

Even if the fixed factor is free (i.e., costs nothing), the rational entrepreneur will stop only at the end of stage 1 (i.e., at point N) where the average product of the variable factor is maximum. At the end point N of stage 1, the producer they will be making maximum use of the fixed factor.

It is thus clear from above that the rational producer will never be found producing in stage 1 and stage 3. Stage 1 and 3 may, therefore, be called stages of economic absurdity or economic non-sense. The stages 1 and 3 represent non-economic regions in production function.

A rational producer will always seek to produce in stage 2 where both the marginal product and average product of the variable factor are diminishing. At which particular point in this stage, the producer will decide to produce depends upon the prices of factors. The stage 2 represents the range of rational production decisions.

We have seen above how output varies as the factor proportions are altered at any given moment. We have also noticed that this input-output relation can be divided into three stages. Now, the question arises as to what causes increasing marginal returns to the variable factor in the beginning, diminishing marginal returns later and negative marginal returns to the variable factor ultimately.

Causes of Initial Increasing marginal Returns to a Factor:

In the beginning, the quantity of the fixed factor is abundant relative to the quantity of the variable factor. Therefore, when more and more units of a variable factor are added to the constant quantity of the fixed factor, the fixed factor is more intensively and effectively utilised.

This causes the production to increase at a rapid rate. When, in the beginning the variable factor is relatively smaller in quantity, some amount of the fixed factor may remain unutilised and therefore when the variable factor is increased fuller utilisation of the fixed factor becomes possible with the result that increasing returns are obtained.

The question arises as to why the fixed factor is not initially taken in an appropriate quantity which suits the available quantity of the variable factor. Answer to this question is provided by the fact that generally those factors are taken as fixed which are indivisible. Indivisibility of a factor means that due to technological requirements a minimum amount of that factor must be employed whatever the level of output.

Thus, as more units of variable factor are employed to work with an indivisible fixed factor, output greatly increases in the beginning due to fuller and more effective utilisation of the latter. Thus, we see that it is the indivisibility of some factors which causes increasing returns to the variable factor in the beginning.

The second reason why we get increasing returns to the variable factor in the initial stage is that as more units of the variable factor are employed the efficiency of the variable factor itself increases. This is because when there is a sufficient quantity of the variable factor, it becomes possible to introduce specialisation or division of labour which results in higher productivity. The greater the quantity of the variable factor, the greater the scope of specialisation and hence the greater will be the level of its productivity or efficiency.

Causes of Diminishing marginal Returns to a Factor

The stage of diminishing marginal returns in the production function with one factor variable is the most important. The question arises as to why we get diminishing marginal returns after a certain amount of the variable factor has been added to a fixed quantity of the other factor.

As explained above, increasing returns to a variable factor occur initially primarily because of the more effective and fuller use of the fixed factor becomes possible as more units of the

variable factor are employed to work with it. Once the point is reached at which the amount of the variable factor is sufficient to ensure the efficient utilisation of the fixed factor, then further increases in the variable factor will cause marginal and average products of a variable factor to decline because the fixed factor then becomes inadequate relative to the quantity of the variable factor.

In other words, the contributions to the production made by the variable factor after a point become less and less because the additional units of the variable factor have less and less of the fixed factor to work with. The production is the result of the co-operation of various factors aiding each other. Now, how much aid one factor provides to the others depends upon how much there is of it.

Eventually, the fixed factor is abundant relative to the number of the variable factor and the former provides much aid to the latter. Eventually, the fixed factor becomes more and more scarce in relation to the variable factor so that as the units of the variable factor are increased they receive less and less aid from the fixed factor. As a result, the marginal and average products of the variable factor decline ultimately.

The phenomenon of diminishing marginal returns, like that of increasing marginal returns, rests upon the indivisibility of the fixed factor. As explained above, the important reason for increasing returns to a factor in the beginning is the fact that the fixed factor is indivisible which has to be employed whether the output to be produced is small or large.

When the indivisible fixed factor is not being fully used, successive increases in a variable factor add more to output since fuller and more efficient use is made of the indivisible fixed factor. But there is generally a limit to the range of employment of the variable factor over which its marginal and average products will increase.

There will usually be a level of employment of the Variable factor at which indivisible fixed factor is being as fully and efficiently used as possible. It will happen when the variable factor

has increased to such an amount that the fixed indivisible factor is being used in the "best or optimum proportion" with the variable factor.

Once the optimum proportion is disturbed by further increases in the variable factor, returns to a variable factor (i.e., marginal product and average product) will diminish primarily because the indivisible factor is being used too intensively, or in other words, the fixed factor is being used in non-optimal proportion with the variable factor.

Just as the marginal product of the variable factor increases in the first stage when better and fuller use of the fixed indivisible factor is being made, so the marginal product of the variable factor diminishes when the fixed indivisible factor is being worked too hard.

If the fixed factor was perfectly divisible, neither the increasing nor the diminish-ing returns to a variable factor would have occurred. If the factors were perfectly divisible, then there would not have been the necessity of taking a large quantity of the fixed factor in the beginning to combine with the varying quantities of the other factor.

In the presence of perfect divisibility, the optimum proportion between the factors could have always been achieved. Perfect divisibility of the factors implies that a small firm with a small machine and one worker would be as efficient as a large firm with a large machine and many workers.

The productivity of the factors would be the same in the two cases. Thus, we see that if the factors were perfectly divisible, then the question of varying factor proportions would not have arisen and hence the phenomena of increasing and diminishing marginal returns to a variable factor would not have occurred. Prof. Bober rightly remarks: "Let divisibility enter through the door, law of variable proportions rushes out through the window."

Joan Robinson goes deeper into the causes of diminishing returns. She holds that the diminish-ing marginal returns occur because the factors of production are imperfect substitutes for one another. As seen above, diminishing returns occur during the

second stage since the fixed factor is now inadequate relatively to the variable factor. Now, a factor which is scarce in supply is taken as fixed.

When there is a scarce factor, quantity of that factor cannot be increased in accordance with the varying quantities of the other factors, which, after the optimum proportion of factors is achieved, results in diminishing returns.

If now some factors were available which perfect substitute of the scarce fixed factor was, then the paucity of the scarce fixed factor during the second stage would have been made up by the increase in supply of its perfect substitute with the result that output could be expanded without diminishing returns.

Thus, even if one of the variable factors which we add to the fixed factor were perfect substitute of the fixed factor, then when, in the second stage, the fixed factor becomes relatively deficient, its deficiency would have been made up the increase in the variable factor which is its perfect substitute.

Thus, Joan Robinson says, "What the Law of Diminishing Returns really states is that there is a limit to the extent to which one factor of production can be substituted for another, or, in other words, that the elasticity of substitution between factor is not infinite.

If this were not true, it would be possible, when one factor of production is fixed in amount and the rest are in perfectly elastic supply, to produce part of the output with the aid of the fixed factor, and then, when the optimum proportion between this and other factors was attained, to substitute some other factor for it and to increase output at constant cost." We, therefore, see that diminishing returns operate because the elasticity of substitution between factors is not infinite.

Explanation of Negative Marginal Returns to a Factor

As the amount of a variable factor continues to be increased to a fixed quantity of the other factor, a stage is reached when the total product declines and the marginal product of the variable factor becomes negative.

This phenomenon of negative marginal returns to the variable factor in stage 3 is due to the fact that the number of the variable factor becomes too excessive relative to the fixed factor so that they obstruct each other with the result that the total output falls instead of rising. Besides, too large a number of the variable factor also impairs the efficiency of the fixed factor. The proverb “too many cooks spoil the broth” aptly applies to this situation. In such a situation, a reduction in the units of the variable factor will increase the total output.

LAWS OF RETURNS TO SCALE

The law of returns to scale explains the proportional change in output with respect to proportional change in inputs. In other words, the law of returns to scale states when there are a proportionate change in the amounts of inputs, the behavior of output also changes.

The degree of change in output varies with change in the amount of inputs. For example, an output may change by a large proportion, same proportion, or small proportion with respect to change in input. On the basis of these possibilities, law of returns can be classified into three categories:

Economies & Diseconomies of Scale

These occur when mass producing a good results in lower average cost. Economies of scale occur within an firm (internal) or within an industry (external).

Average costs fall per unit – $\text{Average costs per unit} = \frac{\text{total costs}}{\text{quantity produced}}$

Internal Economies of Scale - As a business grows in scale, its costs will fall due to internal economies of scale. An ability to produce units of output more cheaply.

External Economies of Scale - Are those shared by a number of businesses in the same industry in a particular area.

Types of Internal Economies of Scale

Below are types of internal economy of scale

Production / Technical Economies

- Larger firms can use computers / technology to replace workers on a production line
- Mass production lowers cost per unit
- Large scale producers can employ techniques that are unable to be used by a small scale producer.
- Able to transport bulk materials.

Purchasing / Marketing Economies

- Advertising costs can be spread across products
- Large businesses can employ specialist staff
- Bulk buying - if you buy more unit cost falls

Financial Economies

- Larger firms have better lending terms and lower rates of interest
- Easier for large firms to raise capital.
- Risk is spread over more products.
- Greater potential finance from retained profits.
- Administration costs can be divided amongst more products

Managerial Economies

- More specialised management can be employed, this increases the efficiency of the business decreasing the costs

Risk-bearing Economies

- Large firms are more likely to take risks with new products as they have more products to spread the risk over

External Economies of Scale

These are advantages gained for the whole industry, not just for individual businesses.

For Example:

- As businesses grow within an area, specialist skills begin to develop.
- Skilled labour in the area – local colleges may begin to run specialist courses.

- Being close to other similar businesses who can work together with each other.
- Having specialist supplies and support services nearby.
- Reputation

Diseconomies of Scale

Occur when firms become too large or inefficient. Average costs per unit start to rise. Below are the types of diseconomy of scale and some examples

Communication

- When firms grow there can be problems with communication
- As the number of people in the firm increases it is hard to get the messages to the right people at the right time
- In larger businesses it is often difficult for all staff to know what is happening

Coordination and control problems

- As a business grows control of activities gets harder
- As the firm gets bigger and new parts of the business are set up it is increasingly likely people will be working in different ways and this leads to problems with monitoring

Motivation

- As businesses grow it is harder to make everyone feel as though they belong
- Less contact between senior managers and employees so employees can feel less involved
- Smaller businesses often have a better team environment which is lost when they grow

Economies of Scale & Monopolies

- Economies of scale can lead to the development of monopolies as larger businesses are able to exploit lower unit costs and therefore make more profits

Minimum efficient scale

- Where an increase in the scale of production gives no benefits to a reduction in unit costs

- This is the point where production is sufficient for internal economies of scale to be fully exploited
- Minimum efficient scale is seen as the lowest point on the long run average cost curve
- The MES depends on a number of factors including: Ratio of fixed to variable costs If a natural monopoly exists

Minimum efficient plant size - Where an increase in the scale of production of an individual plant within the industry doesn't result in any unit cost benefits

Economies of Scale & Barriers to Entry

Economies of scale can act as a barrier to entry for firms into a market. This is because economies of scale allow a firm to have a lower cost structure and therefore can decrease prices if a new firm enters the market eventually driving them out

Economies and Diseconomies of Scale

Economies of scale are defined as the cost advantages that an organization can achieve by expanding its production in the long run.

In other words, these are the advantages of large scale production of the organization. The cost advantages are achieved in the form of lower average costs per unit.

It is a long term concept. Economies of scale are achieved when there is an increase in the sales of an organization. As a result, the savings of the organization increases, which further enables the organization to obtain raw materials in bulk. This helps the organization to enjoy discounts. These benefits are called as economies of scale.

The economies of scale are divided into internal economies and external economies discussed as follows:

Internal Economies

Refer to real economies which arise from the expansion of the plant size of the organization. These economies arise from the growth of the organization itself.

The examples of internal economies of scale are as follows:

- a. Technical economies of scale: Occur when organizations invest in the expensive and advanced technology. This helps in lowering and controlling the costs of production of organizations. These economies are enjoyed because of the technical efficiency gained by the organizations. The advanced technology enables an organization to produce a large number of goods in short time. Thus, production costs per unit falls leading to economies of scale.
- b. Marketing economies of scale: Occur when large organizations spread their marketing budget over the large output. The marketing economies of scale are achieved in case of bulk buying, branding, and advertising. For instance, large organizations enjoy benefits on advertising costs as they cover larger audience. On the other hand, small organizations pay equal advertising expenses as large organizations, but do not enjoy such benefits on advertising costs.
- c. Financial economies of scale: Take place when large organizations borrow money at lower rate of interest. These organizations have good credibility in the market. Generally, banks prefer to grant loans to those organizations that have strong foothold in the market and have good repaying capacity.
- d. Managerial economies of scale: Occur when large organizations employ specialized workers for performing different tasks. These workers are experts in their fields and use their knowledge and experience to maximize the profits of the organization. For instance, in an organization, accounts and research department are created and managed by experienced individuals, SO that all costs and profits of the organization can be estimated properly.
- e. Commercial economies: Refer to economies in which organizations enjoy benefits of buying raw materials and selling of finished goods at lower cost. Large organizations buy raw materials in bulk; therefore, enjoy benefits in transportation charges, easy credit from banks, and prompt delivery of products to customers.

External economies

Occur outside the organization. These economies occur within the industries which benefit organizations. When an industry expands, organizations may benefit from better transportation network, infrastructure, and other facilities. This helps in decreasing the cost of an organization.

Some of the examples of external economies of scale are discussed as follows:

- a. Economies of Concentration: Refer to economies that arise from the availability of skilled labor, better credit, and transportation facilities.
- b. Economies of Information: Imply advantages that are derived from publication related to trade and business. The central research institutions are the source of information for organizations.
- c. Economies of Disintegration: Refer to the economies that arise when organizations split their processes into different processes.

Diseconomies of scale occur when the long run average costs of the organization increases. It may happen when an organization grows excessively large. In other words, the diseconomies of scale cause larger organizations to produce goods and services at increased costs.

There are two types of diseconomies of scale, namely, internal diseconomies and external diseconomies, discussed as follows:

- i. Internal diseconomies of scale: Refer to diseconomies that raise the cost of production of an organization. The main factors that influence the cost of production of an organization include the lack of decision, supervision, and technical difficulties.
- ii. External diseconomies of scale: Refer to diseconomies that limit the expansion of an organization or industry. The factors that act as restraint to expansion include increased cost of production, scarcity of raw materials, and low supply of skilled laborer.

There are a number of causes for diseconomies of scale.

Some of the causes which lead to diseconomies of scale are as follows:

- i. **Poor Communication:** Act as a major reason for diseconomies of scale. If production goals and objectives of an organization are not properly communicated to employees within the organization, it may lead to overproduction or production. This may lead to diseconomies of scale.

Apart from this, if the communication process of the organization is not strong then the employees would not get adequate feedback. As a result, there would be less face-to-face interaction among employees- thus the production process would be affected.

- ii. **Lack of Motivation:** Leads to fall in productivity levels. In case of a large organization, workers may feel isolated and are less appreciated for their work, thus their motivation diminishes. Due to poor communication network, it is harder for employers to interact with the employees and build a sense of belongingness. This leads to fall in the productivity levels of output owing to lack of motivation. This further leads to increase in costs of the organization.
- iii. **Loss of Control:** Acts as the main problem of large organizations. Monitoring and controlling the work of every employee in a large organization becomes impossible and costly. It is harder to make out that all the employees of an organization are working towards the same goal. It becomes difficult for managers to supervise the subordinates in large organizations.
- iv. **Cannibalization:** Implies a situation when an organization faces competition from its own product. A small organization faces competition from products of other organizations, whereas sometimes large organizations find that their own products are competing with each other.

Internal Economies and Diseconomies of Scale: Meaning and Types

As a firm changes its scale of operation, its average costs are likely to change. The usual U-shaped LRAC curve. Average costs fall at first, reach an optimum point and then rise.

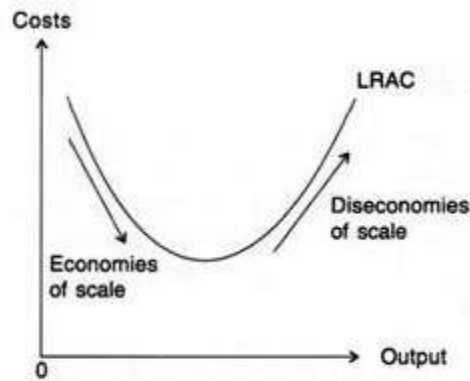


Fig. 1 Internal economies and diseconomies of scale

In very capital intensive industries, such as oil refining, long run average costs may fall over a considerable range of output.

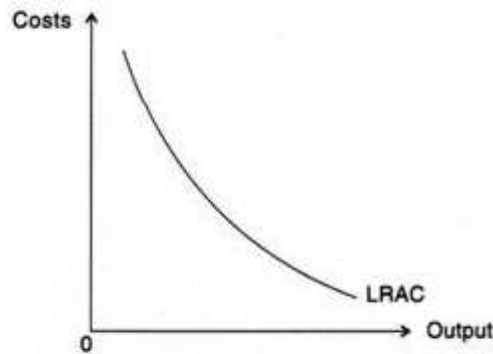


Fig. 2 Downward sloping LRAC curve

In other cases, average costs may fall relatively quickly to their lowest point (the minimum efficient scale) and then remain constant over a large range of output.

This would give an L-shaped LRAC curve.

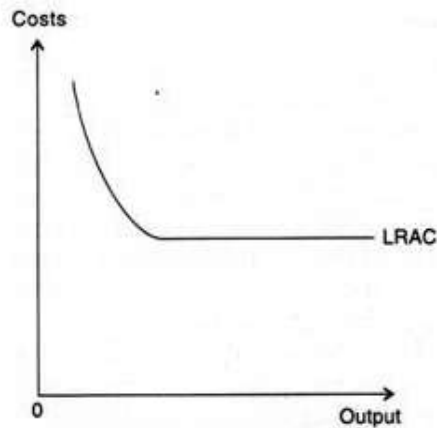


Fig. 3 L-shaped LRAC curve

External economies and diseconomies of scale have a different effect on a firm's LRAC curve.

In the case of external economies of scale, a firm's average costs will be reduced not by the changes in its own output but by the changes in the industry's output. How external economies of scale result in a downward shift of a firm's LRAC curve.

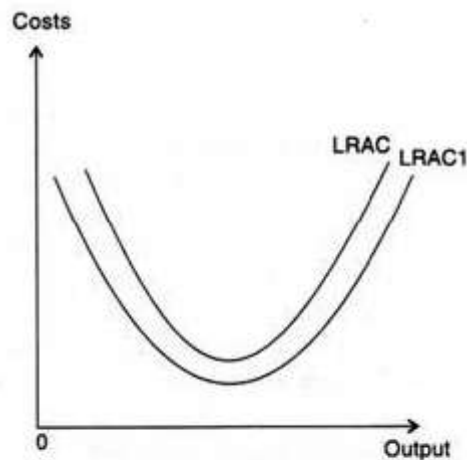


Fig. 4 The effect of external economies of scale

In contrast, external diseconomies of scale will raise a firm's LRAC curve at each and every level of output.

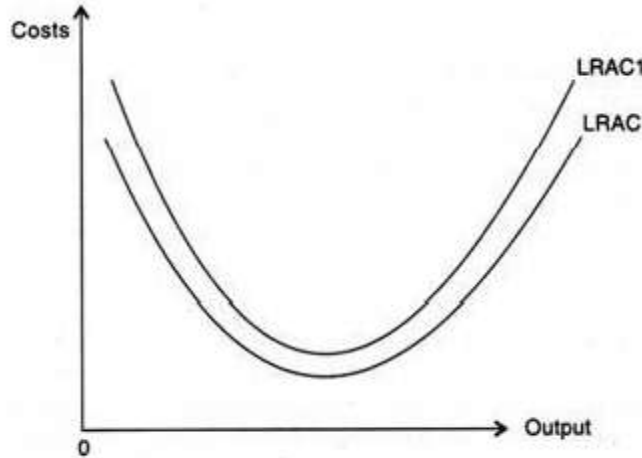


Fig. 5 The effect of external diseconomies of scale

Types of internal economies of scale

As a firm increases its scale of operation, there are a number of reasons responsible for a decline in its average cost. These include:

- i. **Buying economies:** These are probably the best known type. Large firms that buy raw materials in bulk and place large orders for capital equipment usually receive discount. This means that they pay less for each item purchased. They may also receive better treatment than small firms in terms of quality of the raw materials and capital equipment sold and the speed of delivery. This is because the suppliers will be anxious to keep such large customers.
- ii. **Selling economies:** The total cost of processing orders, packing the goods and transporting them does not rise in line with the number of orders. For instance, it costs less than twice as much to send 10,000 washing machines to customers than it does to send 5,000 washing machines. A lorry that can transport 40 washing machines does not

cost four times as much to operate as four vans which can carry 10 washing machines each.

A large volume of output can also reduce advertising costs. The total cost of an advertising campaign can be spread over more units and, again, discounts may be secured. A whole page advertisement in a newspaper or magazine is usually less than twice the cost of a half page advertised. Together, buying and selling economies of scale are sometimes referred to as marketing economies.

- iii. **Managerial economies:** Large firms can afford to employ specialist staff as they can spread their pay over a high number of units. Employing specialist buyers, accountants, human resource managers and designers can increase the firm's efficiency, reduce costs of production and raise demand and revenue. Large firms can also engage in division of labour amongst their other staff. For example, car workers specialize in a particular aspect of the production process.
- iv. **Financial economies:** Large firms usually find it easier and cheaper, to raise finance. Banks tend to be more willing to lend to large firms because such firms are well known and have valuable assets to offer as collateral. Banks often charge large borrowers less, per \$ borrowed, in order to attract them and because they know that the administrative costs of operating and processing large loans are not significantly higher than the costs of dealing with small loans.

Large firms can also raise finance through selling shares which is not an available option for sole traders and partnerships. Public limited companies can sell to the general public. The larger and better known the companies are, the more willing people are to buy their shares.
- v. **Technical economies:** The larger the output of a firm, the more viable it becomes to use large, technologically advanced machinery. Such machinery is likely to be efficient, producing output at a lower average cost than small firms.

- vi. Research and development economies: A large firm can have a research and development department, since running such a department can reduce average costs by developing more efficient methods of production and raise total revenue by developing new products.
- vii. Risk bearing economies: Larger firms usually produce a range of products. This enables them to spread the risks of trading. If the profitability of one of the products it produces falls, it can shift its resources to the production of more profitable products.

Internal diseconomies of scale

Growing beyond a certain output can cause a firm's average costs to rise. This is because a firm may encounter a number of problems including:

- i. Difficulties controlling the firm: It can be hard for those managing a large firm to supervise everything that is happening in the business. Management becomes more complex. A number of layers of management may be needed and there may be a need for more meetings. This can increase administrative costs and make the firm slower in responding to changes in market conditions.
- ii. Communication problems: It can be difficult to ensure that everyone in a large firm have full knowledge about their duties and available opportunities (like training etc.). Also, they may not get the opportunity to effectively communicate their views and ideas to the management team.
- iii. Poor industrial relations: Large firms may be at a greater risk from a lack of motivation of workers, strikes and other industrial action. This is because workers may have less sense of belonging, longer time may be required to solve problems and more conflicts may arise due to the presence of diverse opinions.

Increasing Returns to Scale

If the proportional change in the output of an organization is greater than the proportional change in inputs, the production is

said to reflect increasing returns to scale. For example, to produce a particular product, if the quantity of inputs is doubled and the increase in output is more than double, it is said to be an increasing returns to scale. When there is an increase in the scale of production, the average cost per unit produced is lower. This is because at this stage an organization enjoys high economies of scale. Now, the combination of inputs has reached to $2K+2L$ from $1K+1L$. However, the output has increased from 10 to 25 (150% increase), which is more than double.

Similarly, when input changes from $2K+2L$ to $3K+3L$, then output changes from 25 to 50 (100% increase), which is greater than change in input. This shows increasing returns to scale. There are a number of factors responsible for increasing returns to scale.

Some of the factors are as follows:

Technical and managerial indivisibility

Implies that there are certain inputs, such as machines and human resource, used for the production process are available in a fixed amount. These inputs cannot be divided to suit different level of production. For example, an organization cannot use the half of the turbine for small scale of production. Similarly, the organization cannot use half of a manager to achieve small scale of production. Due to this technical and managerial indivisibility, an organization needs to employ the minimum quantity of machines and managers even in case the level of production is much less than their capacity of producing output. Therefore, when there is increase in inputs, there is exponential increase in the level of output.

Specialization

Implies that high degree of specialization of man and machinery helps in increasing the scale of production. The use of specialized labor and machinery helps in increasing the productivity of labor and capital per unit. This results in increasing returns to scale.

Concept of Dimensions

Refers to the relation of increasing returns to scale to the

concept of dimensions. According to the concept of dimensions, if the length and breadth of a room increases, then its area gets more than doubled. For example, length of a room increases from 15 to 30 and breadth increases from 10 to 20. This implies that length and breadth of room get doubled. In such a case, the area of room increases from 150 (15×10) to 600 (30×20), which is more than doubled.

Constant Returns to Scale

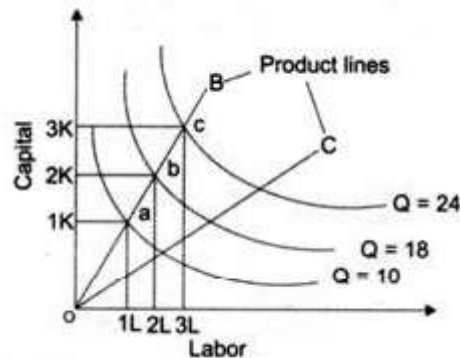
The production is said to generate constant returns to scale when the proportionate change in input is equal to the proportionate change in output. For example, when inputs are doubled, so output should also be doubled, then it is a case of constant returns to scale.

Similarly, when input changes from $2K + 2L$ to $3K + 3L$, then output changes from 20 to 30, which is equal to the change in input. This shows constant returns to scale. In constant returns to scale, inputs are divisible and production function is homogeneous.

Diminishing Returns to Scale

Diminishing returns to scale refers to a situation when the proportionate change in output is less than the proportionate change in input. For example, when capital and labor is doubled but the output generated is less than doubled, the returns to scale would be termed as diminishing returns to scale.

The diminishing returns to scale:



The combination of labor and capital moves from point a to point b, it indicates that input is doubled. At point a, the combination of input is $1K+1L$ and at point b, the combination becomes $2K+2L$.

However, the output has increased from 10 to 18, which is less than change in the amount of input. Similarly, when input changes from $2K+2L$ to $3K + 3L$, then output changes from 18 to 24, which is less than change in input. This shows the diminishing returns to scale.

Diminishing returns to scale is due to diseconomies of scale, which arises because of the managerial inefficiency. Generally, managerial inefficiency takes place in large-scale organizations. Another cause of diminishing returns to scale is limited natural resources. For example, a coal mining organization can increase the number of mining plants, but cannot increase output due to limited coal reserves.

Law of Returns to Scale : Definition, Explanation and Its Types

In the long run all factors of production are variable. No factor is fixed. Accordingly, the scale of production can be changed by changing the quantity of all factors of production.

Definition

“The term returns to scale refers to the changes in output as all factors change by the same proportion.” Koutsoyiannis

“Returns to scale relates to the behaviour of total output as all inputs are varied and is a long run concept”. Leibhafskey

Returns to scale are of the following three types:

1. Increasing Returns to scale.
2. Constant Returns to Scale
3. Diminishing Returns to Scale

Explanation

In the long run, output can be increased by increasing all factors in the same proportion. Generally, laws of returns to scale refer to an increase in output due to increase in all factors in the

same proportion. Such an increase is called returns to scale. Suppose, initially production function is as follows:

$$P = f(L, K)$$

Now, if both the factors of production i.e., labour and capital are increased in same proportion i.e., x , product function will be rewritten as.

$$P_1 = f(xL, xK)$$

1. If P_1 increases in the same proportion as the increase in factors of production i.e., $\frac{P_1}{P} = x$, it will be constant returns to scale.

2. If P_1 increases less than proportionate increase in the factors of production i.e., $\frac{P_1}{P} < x$, it will be diminishing returns to scale.

3. If P_1 increases more than proportionate increase in the factors of production, i.e., $\frac{P_1}{P} > x$, it will be increasing returns to scale. Returns to scale can be shown with the help of table 8.

Table 8. Showing different stages of return to scale

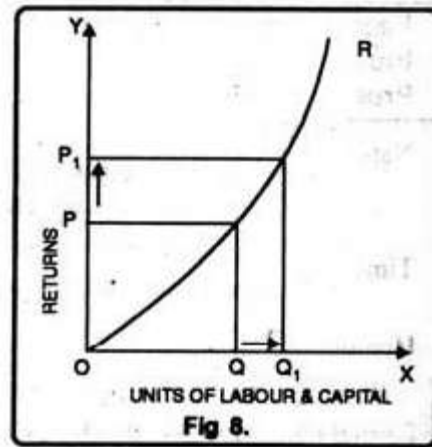
Units of Labour	Units of capital	%age increase in Labour & Capital	Total Product	%age increase in TP	Returns to scale
1	3	—	10	—	Increasing
2	9	100%	30	200%	
3	9	50%	60	100%	
4	12	33%	80	33%	Constant
5	15	25%	100	25%	
6	18	20%	120	10%	Decreasing
7	21	16.6%	130	8.3%	

Increasing Returns to Scale

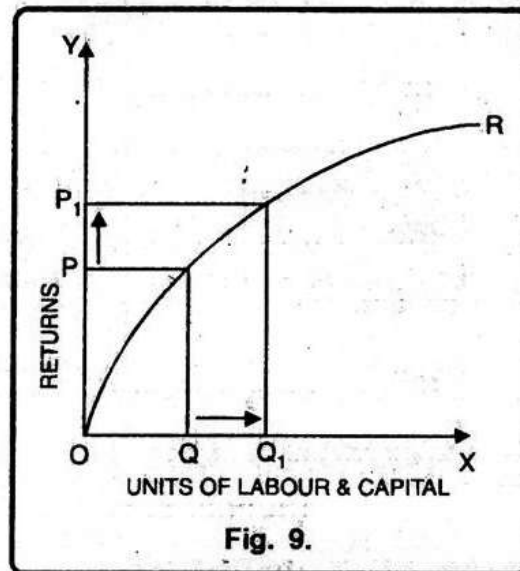
Increasing returns to scale or diminishing cost refers to a situation when all factors of production are increased, output increases at a higher rate.

It means if all inputs are doubled, output will also increase at the faster rate than double. Hence, it is said to be increasing returns to scale. This increase is due to many reasons like division external economies of scale. Increasing returns to scale.

OX axis represents increase in labour and capital while OY axis shows increase in output. When labour and capital increases from Q to Q_1 , output also increases from P to P_1 which is higher than the factors of production i.e. labour and capital.



Diminishing Returns to Scale



Diminishing returns or increasing costs refer to that production situation, where if all the factors of production are increased in a given proportion, output increases in a smaller proportion. It means, if inputs are doubled, output will be less than doubled.

If 20 percent increase in labour and capital is followed by 10 percent increase in output, then it is an instance of diminishing returns to scale.

The main cause of the operation of diminishing returns to scale is that internal and external economies are less than internal and external diseconomies.

The diminishing returns to scale has been shown. On OX axis, labour and capital are given while on OY axis, output. When factors of production increase from Q to Q_1 (more quantity) but as a result increase in output, i.e. P to P_1 is less. We see that increase in factors of production is more and increase in production is comparatively less, thus diminishing returns to scale apply.

Constant Returns to Scale

Constant returns to scale or constant cost refers to the production situation in which output increases exactly in the same proportion in which factors of production are increased. In simple terms, if factors of production are doubled output will also be doubled.

In this case internal and external economies are exactly equal to internal and external diseconomies. This situation arises when after reaching a certain level of production, economies of scale are balanced by diseconomies of scale. This is known as homogeneous production function. Cobb-Douglas linear homogenous production function is a good example of this kind. We see that increase in factors of production i.e. labour and capital are equal to the proportion of output increase. Therefore, the result is constant returns to scale.

ISO-QUANTS: MEANING AND PROPERTIES

Isoquants

There are many ways of producing a given level of output. You can use a lot of labour with a minimal amount of capital, or you could invest heavily in capital equipment that requires a minimal amount of labour to operate, or any combination in between. For most goods, there are more than just two inputs.

For example in agriculture, the amount of land, water, and fertilizer can all be varied to produce different amounts of a crop. An isoquant, in the two input case, is a curve that shows all the ways of combining two inputs so as to produce a given level of output. In the three input case it will be a surface. *Iso* is Latin for equal and *quant* is short for quantity.

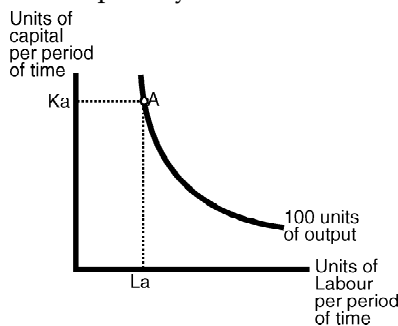


Fig. Isoquant

Movement along an isoquant depicts a constant rate of output, but a changing input ratio. A unique isoquant can be constructed for every level of output, and a family of isoquants can be created to represent various output levels. Isoquants further from the origin represent greater amounts of output. Isoquants are usually considered to be *everywhere dense*, meaning an infinite number of them could be plotted in any two input space.

A typical isoquant is illustrated in the diagram to the right. At point A in the diagram K_a units of capital are combined with L_a units of labour to produce 100 units of output. It is downward sloping, convex to the origin, and non-intersecting (additional isoquants, not shown, would be drawn parallel to this one).

A complete isoquant is actually a closed curve, but only the “down sloping to the right” portion makes economic sense. The upward sloping parts of isoquants, for example, indicate that that level of output could be produced by less of both inputs so this chapter is of little interest to decision makers. The economic section of the isoquants is defined by a pair of lines called *ridge lines*. The “downward to the right” slope of the economic region of an isoquant is due to the possibility of substituting one input for

another in the production process while keeping the level of output constant.

The Marginal rate of Technical Substitution

Isoquants are typically convex to the origin reflecting the fact that the two factors are substitutable for each other at varying rates. This rate of substitutability is called the “marginal rate of technical substitution” (MRTS) or occasionally the “marginal rate of substitution in production”. It measures the reduction in one input per unit increase in the other input that is just sufficient to maintain a constant level of production. For example, the marginal rate of substitution of labour for capital gives the amount of capital that can be replaced by one unit of labour while keeping output unchanged.

To move from point A to point B in the diagram, the amount of capital is reduced from K_a to K_b while the amount of labour is increased only from L_a to L_b .

To move from point C to point D, the amount of capital is reduced from K_c to K_d while the amount of labour is increased from L_a to L_b . The marginal rate of technical substitution of labour for capital is equivalent to the absolute slope of the isoquant at that point (change in capital divided by change in labour). It is equal to 0 where the isoquant becomes horizontal, and equal to infinity where it becomes vertical.

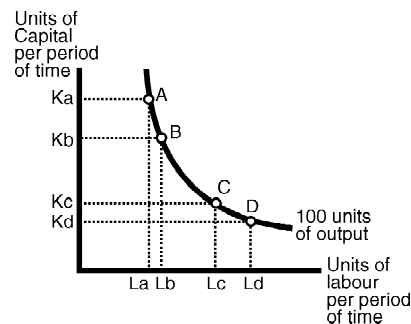


Fig. Marginal Rate of Technical Substitution

The opposite is true when going in the other direction (from D to C to B to A). In this case we are looking at the marginal

rate of technical substitution capital for labour (which is the reciprocal of the marginal rate of technical substitution labour for capital).

It can also be shown that the marginal rate of substitution labour for capital, is equal to the marginal physical product of labour divided by the marginal physical product of capital.

In the unusual case of two inputs that are perfect substitutes for each other in production, the isoquant would be linear (linear, a straight line, with a function $y = a - bx$). If, on the other hand, there is only one production process available, factor proportions would be fixed, and these zero-substitutability isoquants would be shown as horizontal or vertical lines.

LEAST COST COMIBINATION

Combination

These types are not mutually exclusive. Thus a company may vary pricing by location, but then offer bulk discounts as well. Airlines use several different types of price discrimination, including:

- Bulk discounts to wholesalers, consolidators, and tour operators
- Incentive discounts for higher sales volumes to travel agents and corporate buyers
- Seasonal discounts, incentive discounts, and even general prices that vary by location. The price of a flight from say, Singapore to Beijing can vary widely if one buys the ticket in Singapore compared to Beijing (or New York or Tokyo or elsewhere).
- Discounted tickets requiring advance purchase and/or Saturday stays. Both restrictions have the effect of excluding business travelers, who typically travel during the workweek and arrange trips on shorter notice.

Cost Terms and Concepts

Cost can be defined as 'a resource sacrificed or foregone to achieve a specific objective'.

Accountants define costs in monetary terms and while we will focus on monetary costs, readers should recognize that there are not only non-financial measures of performance but also human, social and environmental costs. For example, making employees redundant causes family problems (a human cost) and transfers to society the obligation to pay social security benefits (a social cost).

Pollution causes long-term environmental costs that are also transferred to society. These are as important as (and perhaps more important than) financial costs, but they are not recorded by accounting systems. The exclusion of human, social and environmental costs is a significant limitation of accounting.

For planning, decision-making and control purposes, cost is typically defined in relation to a cost object, which is anything for which a measurement of costs is required. While the cost object is often an output – a product or service – it may also be a resource (an input to the production process), a process of converting resources into outputs or an area of responsibility (a department or cost centre) within the organisation. Examples of inputs are materials, labour, rent, marketing expenses etc. Examples of processes are purchasing, customer order processing, order fulfilment, despatch etc.

Businesses typically report in relation to line items (the resource inputs) and responsibility centres (departments or cost centres).

This means that decisions requiring cost information on business processes and product/service outputs are difficult, because most accounting systems do not provide adequate information about those cost objects.

For example, in a project-based business, published financial reports do not provide cost and revenue information about each project, but instead report information about salaries, rental, office costs etc.

Businesses may adopt a system of management accounting to provide this information for management purposes, but rarely will this second system reconcile with the external financial reports because the management information system may not follow the

same accounting principles. The requirement to produce financial reports based on line items, rather than cost objects, is a second limitation of accounting as a tool of decision-making.

The notion of cost is also problematic because we need to decide how cost is to be defined. If, as Horngren et al. defined it, cost is a resource sacrificed or forgone, then one of the questions we must ask is whether that definition implies a cash cost or an opportunity cost.

A cash cost is the amount of cash expended (a valuable resource), whereas an opportunity cost is the lost opportunity of not doing something, which may be the loss of time or the loss of a customer, equally valuable resources. If it is the cash cost, is it the historical (past) cost or the future cost with which we should be concerned?

For example, is the cost of an employee:

- the historical, cash cost of salaries and benefits, training, recruitment etc. paid? or
- the future cash cost of salaries and benefits to be paid? or
- the lost opportunity cost of what we could have done with the money had
- we not employed that person, e.g. the benefits that could have resulted from expenditure of the same money on advertising, computer equipment, external consulting services etc.?

Wilson and Chua quoted the economist Jevons, writing in 1871, that past costs were irrelevant to decisions about the future because they are 'gone and lost forever'. This is a difficult question and the problematic nature of calculating costs may have been the source of the comment by Clark that there were 'different costs for different purposes'.

Cost Analysis

COST CONCEPT

Cost overrun is defined as excess of actual cost over budget. Cost overrun is also sometimes called “cost escalation,” “cost increase,” or “budget overrun.” However, cost escalation and increases do not necessarily result in cost overruns if cost escalation is included in the budget. Cost overrun is common in infrastructure, building, and technology projects. One of the most comprehensive studies of cost overrun that exists found that 9 out of 10 projects had overrun, overruns of 50 to 100 percent were common, overrun was found in each of 20 nations and five continents covered by the study, and overrun had been constant for the 70 years for which data were available.

For IT projects, an industry study by the Standish Group (2004) found that average cost overrun was 43 percent, 71 percent of projects were over budget, over time, and under scope, and total waste was estimated at US\$55 billion per year in the US alone.

Spectacular examples of cost overrun are the Sydney Opera House with 1,400 percent, and the Concorde supersonic aeroplane with 1,100 percent.

The cost overrun of Boston’s Big Dig was 275 percent, or US\$11 billion. The cost overrun for the Channel tunnel between the UK and France was 80 percent for construction costs and 140 percent for financing costs.

Three types of explanation of cost overrun exist: technical, psychological, and political-economic. Technical explanations account for cost overrun in terms of imperfect forecasting techniques, inadequate data, etc. Psychological explanations account for overrun in terms of optimism bias with forecasters. Finally, political-economic explanations see overrun as the result of strategic misrepresentation of scope and/or budgets.

All of the explanations above can be considered a form of risk. A project's budgeted costs should always include cost contingency funds to cover risks (other than scope changes imposed on the project). As has been shown in cost engineering research, poor risk analysis and contingency estimating practices account for many project cost overruns.

Numerous studies have found that the greatest cause of cost growth was poorly defined scope at the time that the budget was established. The cost growth (overrun of budget before cost contingency is added) can be predicted by rating the extent of scope definition, even on complex projects with new technology.

Cost overrun is typically calculated in one of two ways. Either as a percentage, namely actual cost minus budgeted cost, in percent of budgeted cost. Or as a ratio, viz. actual cost divided by budgeted cost. For example, if the budget for building a new bridge was \$100 million and the actual cost was \$150 million then the cost overrun may be expressed as 50 percent or by the ratio 1.5.

Concept of Cost

Cost accounting is concerned with cost and therefore is necessary to understand the meaning of term cost in a proper perspective. In general, cost means the amount of expenditure (actual or notional) incurred on, or attributable to a given thing.

However, the term cost cannot be exactly defined. Its interpretation depends upon the following factors:

- The nature of business or industry
- The context in which it is used

In a business where selling and distribution expenses are quite nominal the cost of an article may be calculated without considering

the selling and distribution overheads. At the same time, in a business where the nature of a product requires heavy selling and distribution expenses, the calculation of cost without taking into account the selling and distribution expenses may prove very costly to a business. The cost may be factory cost, office cost, cost of sales and even an item of expense.

For example, prime cost includes expenditure on direct materials, direct labour and direct expenses. Money spent on materials is termed as cost of materials just like money spent on labour is called cost of labour and so on. Thus, the use of term cost without understanding the circumstances can be misleading.

Different costs are found for different purposes. The work-in-progress is valued at factory cost while stock of finished goods is valued at office cost. Numerous other examples can be given to show that the term “cost” does not mean the same thing under all circumstances and for all purposes. Many items of cost of production are handled in an optional manner which may give different costs for the same product or job without going against the accepted principles of cost accounting. Depreciation is one of such items. Its amount varies in accordance with the method of depreciation being used. However, endeavour should be, as far as possible, to obtain an accurate cost of a product or service.

Elements of Cost

Following are the three broad elements of cost:

Material : The substance from which a product is made is known as material. It may be in a raw or a manufactured state. It can be direct as well as indirect.

Direct Material : The material which becomes an integral part of a finished product and which can be conveniently assigned to specific physical unit is termed as direct material. Following are some of the examples of direct material:

- All material or components specifically purchased, produced or requisitioned from stores
- Primary packing material (e.g., carton, wrapping, cardboard, boxes etc.)
- Purchased or partly produced components

Direct material is also described as process material, prime cost material, production material, stores material, constructional material etc.

Indirect Material : The material which is used for purposes ancillary to the business and which cannot be conveniently assigned to specific physical units is termed as indirect material. Consumable stores, oil and waste, printing and stationery material etc. are some of the examples of indirect material. Indirect material may be used in the factory, office or the selling and distribution divisions.

Labour: For conversion of materials into finished goods, human effort is needed and such human effort is called labour. Labour can be direct as well as indirect.

Direct Labour : The labour which actively and directly takes part in the production of a particular commodity is called direct labour. Direct labour costs are, therefore, specifically and conveniently traceable to specific products.

Direct labour can also be described as process labour, productive labour, operating labour, etc.

Indirect Labour : The labour employed for the purpose of carrying out tasks incidental to goods produced or services provided, is indirect labour.

Such labour does not alter the construction, composition or condition of the product. It cannot be practically traced to specific units of output. Wages of storekeepers, foremen, timekeepers, directors' fees, salaries of salesmen etc., are examples of indirect labour costs.

Indirect labour may relate to the factory, the office or the selling and distribution divisions.

Expenses : Expenses may be direct or indirect.

Direct Expenses : These are the expenses that can be directly, conveniently and wholly allocated to specific cost centres or cost units. Examples of such expenses are as follows:

- Hire of some special machinery required for a particular contract
- Cost of defective work incurred in connection with a particular job or contract etc.

Direct expenses are sometimes also described as chargeable expenses.

Indirect Expenses : These are the expenses that cannot be directly, conveniently and wholly allocated to cost centres or cost units. Examples of such expenses are rent, lighting, insurance charges etc.

Overhead : The term overhead includes indirect material, indirect labour and indirect expenses. Thus, all indirect costs are overheads.

A manufacturing organization can broadly be divided into the following three divisions:

- Factory or works, where production is done
- Office and administration, where routine as well as policy matters are decided
- Selling and distribution, where products are sold and finally dispatched to customers.

Overheads may be incurred in a factory or office or selling and distribution divisions. Thus, overheads may be of three types:

Factory Overheads : They include the following things:

- Indirect material used in a factory such as lubricants, oil, consumable stores etc.
- Indirect labour such as gatekeeper, timekeeper, works manager's salary etc.
- Indirect expenses such as factory rent, factory insurance, factory lighting etc.

Office and Administration Overheads : They include the following things:

- Indirect materials used in an office such as printing and stationery material, brooms and dusters etc.
- Indirect labour such as salaries payable to office manager, office accountant, clerks, etc.
- Indirect expenses such as rent, insurance, lighting of the office

Selling and Distribution Overheads : They include the following things:

- Indirect materials used such as packing material, printing and stationery material etc.
- Indirect labour such as salaries of salesmen and sales manager etc.
- Indirect expenses such as rent, insurance, advertising expenses etc.

Elements of Cost :

- Direct material
- Direct labour
- Direct expenses
- Overheads
- Factory overheads
- Selling and distribution overheads
- Office and administration overheads
- Indirect material
- Indirect labour
- Indirect expenses
- Indirect material
- Indirect labour
- Indirect expenses
- Indirect material
- Indirect expenses

Components of Total Cost :

1. Prime Cost : Prime cost consists of costs of direct materials, direct labours and direct expenses. It is also known as basic, first or flat cost.
2. Factory Cost : Factory cost comprises prime cost and, in addition, works or factory overheads that include costs of indirect materials, indirect labours and indirect expenses incurred in a factory. It is also known as works cost, production or manufacturing cost.
3. Office Cost : Office cost is the sum of office and administration overheads and factory cost. This is also termed as administration cost or the total cost of production.

4. Total Cost : Selling and distribution overheads are added to the total cost of production to get total cost or the cost of sales.

Cost Sheet

Cost sheet is a document that provides for the assembly of an estimated detailed cost in respect of cost centres and cost units. It analyzes and classifies in a tabular form the expenses on different items for a particular period. Additional columns may also be provided to show the cost of a particular unit pertaining to each item of expenditure and the total per unit cost. Cost sheet may be prepared on the basis of actual data (historical cost sheet) or on the basis of estimated data (estimated cost sheet), depending on the technique employed and the purpose to be achieved.

Classification of Cost

Cost may be classified into different categories depending upon the purpose of classification. Some of the important categories in which the costs are classified are as follows:

Fixed, Variable and Semi-Variable Costs : The cost which varies directly in proportion with every increase or decrease in the volume of output or production is known as variable cost. Some of its examples are as follows:

- Wages of labourers
- Cost of direct material
- Power.

The cost which does not vary but remains constant within a given period of time and a range of activity inspite of the fluctuations in production is known as fixed cost. Some of its examples are as follows:

- Rent or rates
- Insurance charges
- Management salary.

The cost which does not vary proportionately but simultaneously does not remain stationary at all times is known as semi-variable cost. It can also be named as semi-fixed cost. Some of its examples are as follows:

- Depreciation
- Repairs.

Fixed costs are sometimes referred to as “period costs” and variable costs as “direct costs” in system of direct costing. Fixed costs can be further classified into:

- Committed fixed costs
- Discretionary fixed costs.

Committed fixed costs consist largely of those fixed costs that arise from the possession of plant, equipment and a basic organization structure. For example, once a building is erected and a plant is installed, nothing much can be done to reduce the costs such as depreciation, property taxes, insurance and salaries of the key personnel etc. without impairing an organization's competence to meet the long-term goals.

Discretionary fixed costs are those which are set at fixed amount for specific time periods by the management in budgeting process. These costs directly reflect the top management policies and have no particular relationship with volume of output. These costs can, therefore, be reduced or entirely eliminated as demanded by the circumstances. Examples of such costs are research and development costs, advertising and sales promotion costs, donations, management consulting fees etc. These costs are also termed as managed or programmed costs.

In some circumstances, variable costs are classified into the following:

- Discretionary cost
- Engineered cost.

The term discretionary costs is generally linked with the class of fixed cost. However, in the circumstances where management has predetermined that the organization would spend a certain percentage of its sales for the items like research, donations, sales promotion etc., discretionary costs will be of a variable character.

Engineered variable costs are those variable costs which are directly related to the production or sales level. These costs exist in those circumstances where specific relationship exists between

input and output. For example, in an automobile industry there may be exact specifications as one radiator, two fan belts, one battery etc. would be required for one car. In a case where more than one car is to be produced, various inputs will have to be increased in the direct proportion of the output.

Thus, an increase in discretionary variable costs is due to the authorization of management whereas an increase in engineered variable costs is due to the volume of output or sales.

Product Costs and Period Costs : The costs which are a part of the cost of a product rather than an expense of the period in which they are incurred are called as "product costs." They are included in inventory values. In financial statements, such costs are treated as assets until the goods they are assigned to are sold. They become an expense at that time. These costs may be fixed as well as variable, e.g., cost of raw materials and direct wages, depreciation on plant and equipment etc.

The costs which are not associated with production are called period costs. They are treated as an expense of the period in which they are incurred. They may also be fixed as well as variable. Such costs include general administration costs, salaries salesmen and commission, depreciation on office facilities etc. They are charged against the revenue of the relevant period. Differences between opinions exist regarding whether certain costs should be considered as product or period costs. Some accountants feel that fixed manufacturing costs are more closely related to the passage of time than to the manufacturing of a product. Thus, according to them variable manufacturing costs are product costs whereas fixed manufacturing and other costs are period costs. However, their view does not seem to have been yet widely accepted.

Direct and Indirect Costs : The expenses incurred on material and labour which are economically and easily traceable for a product, service or job are considered as direct costs. In the process of manufacturing or production of articles, materials are purchased, labourers are employed and the wages are paid to them. Certain other expenses are also incurred directly. All of these take an active and direct part in the manufacture of a particular commodity and hence are called direct costs.

The expenses incurred on those items which are not directly chargeable to production are known as indirect costs. For example, salaries of timekeepers, storekeepers and foremen. Also certain expenses incurred for running the administration are the indirect costs. All of these cannot be conveniently allocated to production and hence are called indirect costs.

Decision-Making Costs and Accounting Costs : Decision-making costs are special purpose costs that are applicable only in the situation in which they are compiled. They have no universal application. They need not tie into routine-financial accounts. They do not and should not conform the accounting rules. Accounting costs are compiled primarily from financial statements. They have to be altered before they can be used for decision-making. Moreover, they are historical costs and show what has happened under an existing set of circumstances. Decision-making costs are future costs. They represent what is expected to happen under an assumed set of conditions. For example, accounting costs may show the cost of a product when the operations are manual whereas decision-making cost might be calculated to show the costs when the operations are mechanized.

Relevant and Irrelevant Costs : Relevant costs are those which change by managerial decision. Irrelevant costs are those which do not get affected by the decision. For example, if a manufacturer is planning to close down an unprofitable retail sales shop, this will affect the wages payable to the workers of a shop. This is relevant in this connection since they will disappear on closing down of a shop. But prepaid rent of a shop or unrecovered costs of any equipment which will have to be scrapped are irrelevant costs which should be ignored.

Shutdown and Sunk Costs : A manufacturer or an organization may have to suspend its operations for a period on account of some temporary difficulties, e.g., shortage of raw material, non-availability of requisite labour etc. During this period, though no work is done yet certain fixed costs, such as rent and insurance of buildings, depreciation, maintenance etc., for the entire plant will have to be incurred. Such costs of the idle plant are known as shutdown costs. Sunk costs are historical or past costs. These

are the costs which have been created by a decision that was made in the past and cannot be changed by any decision that will be made in the future. Investments in plant and machinery, buildings etc. are prime examples of such costs. Since sunk costs cannot be altered by decisions made at the later stage, they are irrelevant for decision-making.

An individual may regret for purchasing or constructing an asset but this action could not be avoided by taking any subsequent action. Of course, an asset can be sold and the cost of the asset will be matched against the proceeds from sale of the asset for the purpose of determining gain or loss. The person may decide to continue to own the asset. In this case, the cost of asset will be matched against the revenue realized over its effective life. However, he/she cannot avoid the cost which has already been incurred by him/her for the acquisition of the asset. It is, as a matter of fact, sunk cost for all present and future decisions.

Controllable and Uncontrollable Costs : Controllable costs are those costs which can be influenced by the ratio or a specified member of the undertaking. The costs that cannot be influenced like this are termed as uncontrollable costs.

A factory is usually divided into a number of responsibility centres, each of which is in charge of a specific level of management. The officer incharge of a particular department can control costs only of those matter which come directly under his control, not of other matter. For example, the expenditure incurred by tool room is controlled by the foreman incharge of that section but the share of the tool room expenditure which is apportioned to a machine shop cannot be controlled by the foreman of that shop. Thus, the difference between controllable and uncontrollable costs is only in relation to a particular individual or level of management. The expenditure which is controllable by an individual may be uncontrollable by another individual.

Avoidable or Escapable Costs and Unavoidable or Inescapable Costs : Avoidable costs are those which will be eliminated if a segment of a business (e.g., a product or department) with which they are directly related is discontinued.

Unavoidable costs are those which will not be eliminated with the segment. Such costs are merely reallocated if the segment is discontinued. For example, in case a product is discontinued, the salary of a factory manager or factory rent cannot be eliminated. It will simply mean that certain other products will have to absorb a large amount of such overheads. However, the salary of people attached to a product or the bad debts traceable to a product would be eliminated. Certain costs are partly avoidable and partly unavoidable. For example, closing of one department of a store might result in decrease in delivery expenses but not in their altogether elimination.

It is to be noted that only avoidable costs are relevant for deciding whether to continue or eliminate a segment of a business.

Imputed or Hypothetical Costs : These are the costs which do not involve cash outlay. They are not included in cost accounts but are important for taking into consideration while making management decisions. For example, interest on capital is ignored in cost accounts though it is considered in financial accounts. In case two projects require unequal outlays of cash, the management should take into consideration the capital to judge the relative profitability of the projects.

Out-of-Pocket Costs : Out-of-pocket cost means the present or future cash expenditure regarding a certain decision that will vary depending upon the nature of the decision made. For example, a company has its own trucks for transporting raw materials and finished products from one place to another.

It seeks to replace these trucks by keeping public carriers. In making this decision, of course, the depreciation of the trucks is not to be considered but the management should take into account the present expenditure on fuel, salary to drive\$ and maintenance. Such costs are termed as out-of-pocket costs.

Opportunity Cost : Opportunity cost refers to an advantage in measurable terms that have foregone on account of not using the facilities in the manner originally planned.

For example, if a building is proposed to be utilized for housing a new project plant, the likely revenue which the building could

fetch, if rented out, is the opportunity cost which should be taken into account while evaluating the profitability of the project. Suppose, a manufacturer is confronted with the problem of selecting anyone of the following alternatives:

- a. Selling a semi-finished product at \$. 2 per unit
- b. Introducing it into a further process to make it more refined and valuable.

Alternative (b) will prove to be remunerative only when after paying the cost of further processing, the amount realized by the sale of the product is more than \$. 2 per unit. Also, the revenue of \$. 2 per unit is foregone in case alternative (b) is adopted. The term “opportunity cost” refers to this alternative revenue foregone.

Traceable, Untraceable or Common Costs

The costs that can be easily identified with a department, process or product are termed as traceable costs. For example, the cost of direct material, direct labour etc. The costs that cannot be identified so are termed as untraceable or common costs. In other words, common costs are the costs incurred collectively for a number of cost centres and are to be suitably apportioned for determining the cost of individual cost centres. For example, overheads incurred for a factory as a whole, combined purchase cost for purchasing several materials in one consignment etc.

Joint cost is a kind of common cost. When two or more products are produced out of one material or process, the cost of such material or process is called joint cost. For example, when cottonseeds and cotton fibers are produced from the same material, the cost incurred till the split-off or separation point will be joint costs.

SHORT RUN AND LONG RUN COST CURVES

Analysis of Short Run Cost of Production

Short run is a period of time over which at least one factor must remain fixed. For most of the firms, the fixed resource or factors which cannot be increased to meet the rising demand of the good is capital i.e., plant and machinery. Short run, then, is a period of

time over which output can be changed by adjusting the quantities of resources such as labour, raw material, fuel but the size or scale of the firm remains fixed.

Definition of Long Run

In the long run there is no fixed resource. All the factors of production are variable. The length of the long run differs from industry to industry depending upon the nature of production. For example, a balloon making firm can change the size of firm more quickly than a car manufacturing firm.

Categories/Types of Costs in the Short Run

The total cost of a firm in the short run is divided into two categories (1) Fixed cost and (2) Variable cost. The two types of economic costs are now discussed in brief.

- (1) Total Fixed Cost (TFC): *Total fixed cost* occur only in the short run. Total Fixed cost as the name implies is the cost of the firm's fixed resources, Fixed cost remains the same in the short run regardless of how many units of output are produced. We can say that fixed cost of a firm is that part of total cost which does not vary with changes in output per period of time. Fixed cost is to be incurred even if the output of the firm is zero.

For example, the firm's resources which remain fixed in the short run are building, machinery and even staff employed on contract for work over a particular period.

- (2) Total Variable Cost (TVC): *Total variable cost* as the name signifies is the cost of variable resources of a firm that are used along with the firm's existing fixed resources. Total variable cost is linked with the level of output. When output is zero, variable cost is zero. When output increases, variable cost also increases and it decreases with the decrease in output. So any resource which can be varied to increase or decrease with the rate of output is variable cost of the firm.

For example, wages paid to the labour engaged in production, prices of raw material which a firm incurs on

the production of output are variable costs. A firm can reduce its variable cost by lowering output but it cannot decrease its fixed cost. These expenses remain fixed in the short run. In the long run there are no fixed resources. All resources are variable. Therefore, a firm has no fixed cost in the long run. All long run costs are variable costs.

- (3) Total Cost (TC): Total cost is the sum of fixed cost and variable cost incurred at each level of output. Total cost of production of a firm equals its fixed cost plus its:

Formula:

$$TC = TFC + TVC$$

Where:

TC = Total cost.

TFC = Total fixed cost.

TVC = Total variable cost.

Explanation:

Short run costs of a firm is now explained with the help of a schedule and diagrams.

Schedule:

(in Dollars)

<i>Units of Output (in Hundred)</i>	<i>Total Fixed Cost</i>	<i>Total Variable Cost</i>	<i>Total Cost</i>
0	1000	0	1000
1	1000	60	1060
2	1000	100	1100
3	1000	150	1150
4	1000	200	1200
5	1000	400	1400
6	1000	700	1700
7	1000	1100	2100

The short run cost data of the firm shows that total fixed cost TFC (column 2) remains constant at \$1000/—regardless of the level of output. The column 3 indicates variable cost which is associated with the level of output. Total variable cost is zero when production

is zero. Total variable cost increases with the increase in output. The variable does not increase by the same amount for each increase in output. Initially the variable cost increases by a smaller amount up to 3rd unit of output and after which it increases by larger amounts. Column (4) indicates total cost which is the sum of TFC + TVC. The total cost increases for each level of output. The rise in total cost is more sharp after the 4th level of output. The concepts of costs, i.e., (1) total fixed cost (2) total variable cost and (3) total cost can be illustrated graphically.

Short-run Production Analysis

An analysis of the production decision made by a firm in the short run, with the ultimate goal of explaining the law of supply and the upward-sloping supply curve. The central feature of this short-run production analysis is the law of diminishing marginal returns, which results in the short run when larger amounts of a variable input, like labour, are added to a fixed input, like capital. A contrasting analysis is long-run production analysis.

The analysis of short-run production sets the stage to better understand the supply-side of the market. How producers respond to price depends, in part, on their ability to combine inputs to produce output. This ability is guided by the law of diminishing marginal returns, which states that the productivity of a variable input declines as more is added to a fixed input. If productivity declines, then more of the variable input is needed as the quantity produced increases. This results in an increase in production cost, which means producers need to receive a higher price. The connection between higher price and more production is essence of the law of supply.

Two Runs: Short and Long

The first step in the analysis of short-run production is a distinction between the short run and the long run. This distinction is intertwined with the distinction between fixed and variable inputs.

- Short Run: The short run is a period of time in which at least one input used for production and under the control

of the producer is variable and at least one input is fixed.

- Long Run: The long run is a period of time in which at all inputs used for production and under the control of the producer are variable.

The difference between short run and long run depends on the particular production activity. For some producers, the short run lasts a few days. For others, the short run can last for decades.

Two Inputs: Fixed and Variable

The analysis of short-run production assumes that at least one input in the production process is fixed and at least one is variable. As already noted, the fixed and variable inputs are intertwined with the notion of short run and long run.

- Fixed Input: A fixed input is an input used in production and under the control of the producer that does not change during the time period of analysis (the short run).
- Variable Input: A variable input is an input used in production and under the control of the producer that does change during the time period of analysis (the short run).

The variable input used by most producers is more often than not labour. The fixed input for most production operations is usually capital. The presumption is that the size of a firm's workforce can be adjusted more quickly than the size of the factory or building, the amount of equipment, and other capital.

Note that the phrase "under the control of the producer" is included in the specifications of short run, long run, fixed input, and variable input. The reason is that short-run production analysis is most concerned with how producers adjust the inputs under the control in response to changing prices. Any production activity invariably includes inputs (fixed and variable) that are beyond the control of the producer, including government laws and regulations, social customs and institutions, weather, and the forces of nature. These other variables are certainly worthy of consideration, but are not fundamental to explaining and understanding the basic principles of market supply

Three Returns: Increasing, Decreasing, and Negative

The addition of a variable input (like labour) to a fixed input (like capital) can have one of three basic results. First, production might increase at an increasing rate. Second, production might increase at a decreasing rate. Third, production might actually decrease.

These three alternatives are technically termed increasing marginal returns, decreasing marginal returns, and negative marginal returns.

- **Increasing Marginal Returns:** This occurs if each additional unit of a variable input added to a fixed input causes incremental production to increase. For example, the one worker contributes 10 units of output to production, the next worker contributes another 12 units, and the subsequent worker contributes 14 units. With increasing marginal returns, each worker contributes more to production than the previous worker.
- **Decreasing Marginal Returns:** This occurs if each additional unit of a variable input added to a fixed input causes incremental production to decrease. For example, the one worker contributes 10 units of output to production, the next worker contributes another 8 units, and the subsequent worker contributes only 6 units. With decreasing marginal returns, each worker contributes less to production than the previous worker.
- **Negative Marginal Returns:** This results if the addition of a variable input added to a fixed input actually causes the total production to decline. For example, if 10 workers produce a total of 100 units of output, and 11 workers produce a total of 99 units, then the eleventh worker is said to have negative marginal returns.

Most short-run production involves increasing marginal returns with the addition of the first few units of a variable input. This inevitably gives way to decreasing marginal returns. While negative marginal returns are somewhat rare, they do eventually result if too many units of a variable input are added.

One Law

The inevitability of decreasing marginal returns is captured by the most important economic principle in short-run production analysis—the law of diminishing marginal returns.

- The Law of Diminishing Marginal Returns: This law states that as more and more of a variable input is added to a fixed input in short-run production, then the marginal product (that is, the marginal returns) of the variable input eventually declines.

While most short-run production is likely to see increasing marginal returns, eventually, inevitably, most certainly, decreasing marginal returns occur. The law of diminishing marginal returns means that increased production of a good requires more and more of the variable input. For example, the first 50 units of production can be had with only 5 workers. However, the next 50 units might require an additional 10 workers. With more the variable input needed, the cost of production rises. And as the production cost rises, the price that producers need to receive also increases. Hence, a higher price corresponds with a larger quantity, which is the law of supply. The positive law of supply connection between price and quantity, as such, can be traced to the law of diminishing marginal returns.

Three Product Curves

This graph presents the three “product” curves that form the foundation of short-run production analysis. This particular set of curves depict the hourly production of Waldo’s Super Deluxe TexMex Gargantuan Tacos (with sour cream and jalapeno peppers) for different quantities of labour, the variable input. The fixed input is the building, cooking and preparation equipment, cash register, tables, chairs, and other capital that comprise Waldo’s TexMex Taco World restaurant.

- Total Product Curve: The curve labelled TP is the total product curve, the total number of TexMex Gargantuan Tacos produced per hour for a given amount of labour. If Waldo (the owner of Waldo’s TexMex Taco World) hires more employees, he can expect a greater production of

TexMex Gargantuan Tacos until he reaches peak production at 7 and 8 workers. Click the [TP] button to highlight this curve.

- **Marginal Product Curve:** The MP curve is the marginal product curve, and the one that is key to the study of short-run production. The MP curve indicates how the total production of TexMex Gargantuan Tacos changes when an extra worker is hired. For example, hiring a fifth worker means that Waldo's TexMex Taco World can produce an additional 10 TexMex Gargantuan Tacos per hour. Most important, the marginal product declines after the second worker is hired, which is the law of diminishing marginal returns, the driving principle in the study of short-run production. Click the [MP] button to highlight this curve.

Three Production Stages

Short-run production exhibits three distinct stages reflected by the shapes and slopes of the three product curves — total product, marginal product, and average product.

- **Stage I:** The first stage is increasing marginal returns and is characterized by the increasingly steeper positive slope of the total product curve, the positive slope of the marginal product curve, and the positive slope of the average product curve. Moreover, the marginal product curve reaches a peak at the end of Stage I.
- **Stage II:** The second stage is decreasing marginal returns and is reflected in the positive but flattening slope of the total product curve and the negative slope of the marginal product curve. Moreover, the average product reaches a peak and is equal to marginal product in this stage. The marginal product curve intersects the horizontal quantity axis at the end of Stage II.
- **Stage III:** The third and last stage is negative marginal returns illustrated by the negative value of marginal product and the negative slope of the total product curve. Average product is positive, but the average product curve has a negative slope.

One Step

This analysis of short-run production is but the first step in a brisk walk toward a better understanding market supply. Further steps include the cost of short-run production, especially marginal cost, and the market structure in which a firm operates, such as perfect competition or monopoly.

- **Production Cost:** An understanding of market supply builds on the short-run production analysis and the key role played by the law of diminishing marginal returns. Because the productivity of the variable input decreases, a larger quantity is needed as production increases. This larger quantity, however, entails greater production cost, as reflected in a positively-sloped marginal cost curve.
- **Market Structure:** The market supply also depends on the structure of the market, especially the degree of competition and the resulting market control of each firm. Competitive markets, with limited control over the price, tend to produce output by equating price and marginal cost. Because marginal cost increases with production, so too does price. However, less competitive markets, with greater market control by the participating firms, need not equate price and marginal cost. As such, a higher price might not correspond with a larger quantity.

Analytical Importance of Fixed and Variable Costs

In the time of distinction between fixed cost and variable cost is a matter of degree, it all depends upon the contracts of a firm and the period of time under consideration.

For example, if a firm makes contract with the labour for a certain period, then the firm has to bear the cost of the labour irrespective of the total produce.

Under such conditions, the wages paid to the labour will be classified as fixed cost and not variable cost, as discussed under the heading of variable cost. Secondly, when the period of time is short, the distinction between fixed cost and variable cost can

be made rigid but not in a longer period of time all fixed costs change into variable cost in the long run.

Analysis in Product Development

Every company that has been in operation some time has a specific set of products which is not accidental. Studying the reasons which have affected the present choice of products, updating the sortiment to contemporary circumstances and searching an optimal range of future products is an activity called *strategic design*.

The cornerstone in strategic design is the company's business mission which is a broadly defined, enduring statement of purpose that distinguishes a business from others of its type (according to Ackoff, 1987, 30).

On this basis you can continue planning the direction of your company's next operations by posing to yourself a series of questions like:

1. Where are we now (and how did we get there)?
2. Where are we heading, if present trends continue?
3. Where would we like to be (e.g. a *vision* of what the company should be in 5 years)
4. How do we get there?

A more systematic method of evaluating the strategic position of a business is the SWOT analysis where the company's Strengths, Weaknesses, Opportunities and Threats are identified, mostly in relation to competition.

The most usual goal of strategic design is to find ways to expand the business. The main directions where this can be attempted are, either finding new customers, or developing new products. It is also possible to combine these two alternatives (fig. on the right).

- (A). An existing product is marketed to *new* groups of customers.

In this alternative, product development activities are needed only insofar as adjustments to the old product are necessary to enhance its attraction to the customers.

As a contrast, great efforts have to be directed to market research and marketing.

- (B). A completely new product is created and it will be marketed to the existing customer groups.

A prerequisite is that you possess or are able to create the necessary technological innovations. Moreover, you have to modify the innovations on the basis of market research and bring the results together not forgetting the production capacities of the company.

- A new product for new customers, marked with “?” in the diagram, includes all the research and planning operations of both strategies (A and B) and is thus very laborious.

It was earlier often considered as too risky, and so it can be even today. However, in the modern world where everything changes fast it has some chances of success, and a prosperous enterprise can perhaps take the risks.

The strategy “B” necessitates market research which can be done among the existing customers and is thus easily feasible. As a contrast, in the strategy marked “?” you have to find out the preferences of target customers that you do not yet know. To manage this difficult task, you can put to yourself questions like:

- Is there a group of people that share a newly developed mode of living where a certain new product is used?
- What is the pleasant experience that the typical user will enjoy with the product? Could it be intensified by adding some new capabilities to the product?

Answering especially the latter question demands that the product development team knows or understands the evolving life styles of the target group of customers, preferably includes one or two people from this group, and is capable of innovation. The task can take a lot of time, and is is often advisable not to proceed directly to the product concept phase, but instead initiate it with a separate Design Driver innovation phase.

Defining Target Customers

The decisive thing in product development are the desires of

the customers—but exactly who are your customers? Existing customers are an advantage in some fields of industry where you can once and again sell replacements or improvements of your old products to the same customers. Their names and addresses are a precious asset in the development of new products and in their future marketing, and therefore they should be carefully stored into the customer register of the company. Cars are an example of products, the consumers of which are relatively loyal to the brand name, and in this case even the names and updated addresses of a large number of recent customers are well known to the company, because of the yearly visits that the owners of the cars habitually make to the service workshops.

More common, though, is that your new products must be marketed to not exactly the same individuals or firms that were your customers the year before. Anyway, usually your customers will continually belong to the same socio-economic group. If you think that this is the case, you will want to define this target group of customers in such a way that it becomes possible to approach them for market research, perhaps also inviting some of them to participate in the product development team.

When constructing the target group of customers you usually can benefit from already existing, generally used, defined classifications. Candidates for such are e.g. the human populations of countries, regions and cities, and long-established socio-economic classes: school children, students, young people, those living alone, young couples etc. up to old age pensioners. The unemployed constitute a group of its own nowadays. Language, sex and wealth provide also potential partitions.

Generally used statistical groups are often too large or too vague for creating a meaningful target for your products, so the next step is to delimit your target by segmenting. The normal method is to cross tabulate two or more classifications as in the example on the left. Try to think creatively here, as often the most fruitful segmentation is one that no competitor ever came to think of. Try to define a group whose special needs your company could meet better than your competitors can.

Beside public statistical classifications you can sometimes exploit groups that have originated in the commercial or private sector. It would be interesting to see the customer files of your firm's competitors, but that is a book closed to you. Instead, you might consider such material as member lists of certain associations, or subscriber lists of specialized magazines, which are sometimes available for a fee. When defining the target group you should also think about how you shall later approach these people. Marketing can be possible without the exact addresses of potential customers, but not market research. For the purposes of the latter you can use only such registers which contain both the names and addresses of people, or you must be prepared to do the awkward work of finding them out (for those people which have been selected into the sample to be studied). Some of the addresses may be found in the customer register, if your company has one.

There are also firms who sell assorted addresses of potential customers. If the target population turns out to be very large, it may be advisable to pick out a sample of it.

COST: OUTPUT RELATION, MODERN APPROACH OF COST

Classification of Costs

Classification of cost means, the grouping of costs according to their common characteristics. The important ways of classification of costs are:

- By nature or element: materials, labour, expenses
- By functions: production, selling, distribution, administration, R&D, development,
- By traceability: direct and indirect
- By variability: fixed, variable, semi-variable
- By controllability: controllable, uncontrollable
- By normality: normal, abnormal.

Standard Cost Accounting

In modern cost accounting, the concept of recording historical costs was taken further, by allocating the company's

fixed costs over a given period of time to the items produced during that period, and recording the result as the total cost of production. This allowed the *full cost* of products that were not sold in the period they were produced to be recorded in inventory using a variety of complex accounting methods, which was consistent with the principles of GAAP (Generally Accepted Accounting Principles). It also essentially enabled managers to ignore the fixed costs, and look at the results of each period in relation to the “standard cost” for any given product.

For example: if the railway coach company normally produced 40 coaches per month, and the fixed costs were still \$1000/month, then each coach could be said to incur an overhead of \$25 ($\$1000 / 40$). Adding this to the variable costs of \$300 per coach produced a full cost of \$325 per coach.

This method tended to slightly distort the resulting unit cost, but in mass-production industries that made one product line, and where the fixed costs were relatively low, the distortion was very minor.

For example: if the railway coach company made 100 coaches one month, then the unit cost would become \$310 per coach ($\$300 + (\$1000 / 100)$). If the next month the company made 50 coaches, then the unit cost = \$320 per coach ($\$300 + (\$1000 / 50)$), a relatively minor difference.

An important part of standard cost accounting is a variance analysis, which breaks down the variation between actual cost and standard costs into various components (volume variation, material cost variation, labour cost variation, etc.) so managers can understand *why costs were different from what was planned* and take appropriate action to correct the situation.

Inputs and Outputs of the Function

In the basic production function, inputs are typically capital and labor and output is whatever good the firm produces.

A production function relates the input of factors of production to the output of goods. In the basic production function inputs are

typically capital and labor, though more expansive and complex production functions may include other variables such as land or natural resources. Output may be any consumer good produced by a firm. Cars, clothing, sandwiches, and toys are all examples of output.

Capital refers to the material objects necessary for production. Machinery, factory space, and tools are all types of capital. In the short run, economists assume that the level of capital is fixed – firms can't sell machinery the moment it's no longer needed, nor can they build a new factory and start producing goods there immediately. When looking at the production function in the short run, therefore, capital will be a constant rather than a variable. Although in reality a firm may own the capital that it uses, economists typically refer to the ongoing cost of employing capital as the rental rate because the opportunity cost of employing capital is the income that a firm could receive by renting it out. Thus, the price of capital is the rental rate.

Labor refers to the human work that goes into production. Typically economists assume that labor is a variable factor of production; it can be increased or decreased in the short run in order to produce more or less output. The price of labor is the prevailing wage rate, since wages are the cost of hiring an additional unit of capital.

The marginal product of an input is the amount of output that is gained by using one additional unit of that input. It can be found by taking the derivative of the production function in terms of the relevant input. For example, if the production function is $Q=3K+2L$ (where K represents units of capital and L represents units of labor), then the marginal product of capital is simply three; every additional unit of capital will produce an additional three units of output. Inputs are typically subject to the law of diminishing returns: as the amount of one factor of production increases, after a certain point the marginal product of that factor declines.

Elements of Cost of Production

The following elements are included in the cost of production:
(a) Purchase of raw machinery, (b) Installation of plant and

machinery, (c) Wages of labour, (d) Rent of Building, (e) Interest on capital, (f) Wear and tear of the machinery and building, (g) Advertisement expenses, (h) Insurance charges, (i) Payment of taxes, (j) In the cost of production, the imputed value of the factor of production owned by the firm itself is also added, (k) The normal profit of the entrepreneur is also included In the cost of production.

Normal Profit

By *normal profit* of the entrepreneur is meant in economics the sum of money which is necessary to keep an entrepreneur employed in a business. This remuneration should be equal to the amount which he can earn in some other alternative occupation. If this alternative return is not met, he will leave the enterprise and join alternative line of production.

Types/Classifications of Cost of Production

Prof, Mead in his book, "Economic Analysis and Policy" has classified these costs into three main sections:

- (1) Production Costs: It includes material costs, rent cost, wage cost, interest cost and normal profit of the entrepreneur.
- (2) Selling Costs: It includes transportation, marketing and selling costs.
- (3) Sundry Costs: It includes other costs such as insurance charges, payment of taxes and rate, etc., etc.

Production Function: Meaning, Definitions and Features

Production is the result of co-operation of four factors of production viz., land, labour, capital and organization.

This is evident from the fact that no single commodity can be produced without the help of any one of these four factors of production.

Therefore, the producer combines all the four factors of production in a technical proportion. The aim of the producer is to maximize his profit. For this sake, he decides to maximize the production at minimum cost by means of the best combination of

factors of production. The producer secures the best combination by applying the principles of equi-marginal returns and substitution.

According to the principle of equi-marginal returns, any producer can have maximum production only when the marginal returns of all the factors of production are equal to one another. For instance, when the marginal product of the land is equal to that of labour, capital and organisation, the production becomes maximum.

Meaning of Production Function

In simple words, production function refers to the functional relationship between the quantity of a good produced (output) and factors of production (inputs).

“The production function is purely a technical relation which connects factor inputs and output.” Prof. Koutsoyiannis

Defined production function as “the relation between a firm’s physical production (output) and the material factors of production (inputs).” Prof. Watson

In this way, production function reflects how much output we can expect if we have so much of labour and so much of capital as well as of labour etc. In other words, we can say that production function is an indicator of the physical relationship between the inputs and output of a firm.

The reason behind physical relationship is that money prices do not appear in it. However, here one thing that becomes most important to quote is that like demand function a production function is for a definite period.

It shows the flow of inputs resulting into a flow of output during some time. The production function of a firm depends on the state of technology. With every development in technology the production function of the firm undergoes a change.

The new production function brought about by developing technology displays same inputs and more output or the same output with lesser inputs. Sometimes a new production function of the firm may be adverse as it takes more inputs to produce the

same output. Mathematically, such a basic relationship between inputs and outputs may be expressed as:

$$Q = f(L, C, N)$$

Where Q = Quantity of output

L = Labour

C = Capital

N = Land.

Hence, the level of output (Q), depends on the quantities of different inputs (L, C, N) available to the firm. In the simplest case, where there are only two inputs, labour (L) and capital (C) and one output (Q), the production function becomes.

$$Q = f(L, C)$$

Definitions

"The production function is a technical or engineering relation between input and output. As long as the natural laws of technology remain unchanged, the production function remains unchanged." Prof. L.R. Klein

"Production function is the relationship between inputs of productive services per unit of time and outputs of product per unit of time." Prof. George J. Stigler

"The relationship between inputs and outputs is summarized in what is called the production function. This is a technological relation showing for a given state of technological knowledge how much can be produced with given amounts of inputs." Prof. Richard J. Lipsey

Thus, from the above definitions, we can conclude that production function shows for a given state of technological knowledge, the relation between physical quantities of inputs and outputs achieved per period of time.

Features of Production Function

Following are the main features of production function:

1. Substitutability: The factors of production or inputs are substitutes of one another which make it possible to vary

the total output by changing the quantity of one or a few inputs, while the quantities of all other inputs are held constant. It is the substitutability of the factors of production that gives rise to the laws of variable proportions.

2. Complementarity: The factors of production are also complementary to one another, that is, the two or more inputs are to be used together as nothing will be produced if the quantity of either of the inputs used in the production process is zero.

The principles of returns to scale is another manifestation of complementarity of inputs as it reveals that the quantity of all inputs are to be increased simultaneously in order to attain a higher scale of total output.

3. Specificity: It reveals that the inputs are specific to the production of a particular product. Machines and equipment's, specialized workers and raw materials are a few examples of the specificity of factors of production. The specificity may not be complete as factors may be used for production of other commodities too. This reveals that in the production process none of the factors can be ignored and in some cases ignorance to even slightest extent is not possible if the factors are perfectly specific.

Production involves time; hence, the way the inputs are combined is determined to a large extent by the time period under consideration. The greater the time period, the greater the freedom the producer has to vary the quantities of various inputs used in the production process.

In the production function, variation in total output by varying the quantities of all inputs is possible only in the long run whereas the variation in total output by varying the quantity of single input may be possible even in the short run.

Market Structure and Price Determination

PERFECT COMPETITION

In economics, specifically general equilibrium theory, a perfect market, also known as an atomistic market, is defined by several idealizing conditions, collectively called perfect competition, or atomistic competition. In theoretical models where conditions of perfect competition hold, it has been theoretically demonstrated that a market will reach an equilibrium in which the quantity supplied for every product or service, including labor, equals the quantity demanded at the current price. This equilibrium would be a Pareto optimum.

Perfect competition provides both allocative efficiency and productive efficiency:

- Such markets are allocatively efficient, as output will always occur where marginal cost is equal to average revenue i.e. price ($MC = AR$). In perfect competition, any profit-maximizing producer faces a market price equal to its marginal cost ($P = MC$). This implies that a factor's price equals the factor's marginal revenue product. It allows for derivation of the supply curve on which the neoclassical approach is based. This is also the reason why "a monopoly does not have a supply curve". The abandonment of price taking creates considerable difficulties for the demonstration of a general equilibrium except under other,

very specific conditions such as that of monopolistic competition.

- In the short-run, perfectly competitive markets are not necessarily productively efficient as output will not always occur where marginal cost is equal to average cost ($MC = AC$). However, in long-run, productive efficiency occurs as new firms enter the industry. Competition reduces price and cost to the minimum of the long run average costs. At this point, price equals both the marginal cost and the average total cost for each good ($P = MC = AC$).

The theory of perfect competition has its roots in late-19th century economic thought. Léon Walras gave the first rigorous definition of perfect competition and derived some of its main results. In the 1950s, the theory was further formalized by Kenneth Arrow and Gérard Debreu. Real markets are never perfect. Those economists who believe in perfect competition as a useful approximation to real markets may classify those as ranging from close-to-perfect to very imperfect. Share and foreign exchange markets are commonly said to be the most similar to the perfect market. The real estate market is an example of a very imperfect market. In such markets, the theory of the second best proves that if one optimality condition in an economic model cannot be satisfied, it is possible that the next-best solution involves changing other variables away from the values that would otherwise be optimal.

Equilibrium in perfect competition

Equilibrium in perfect competition is the point where market demands will be equal to market supply. A firm's price will be determined at this point. In the short run, equilibrium will be affected by demand. In the long run, both demand and supply of a product will affect the equilibrium in perfect competition. A firm will receive only normal profit in the long run at the equilibrium point.

Characteristics of Perfect Competition

The following are the conditions for the existence of perfect competition:

Large Number of Buyers and Sellers

The first condition is that the number of buyers and sellers must be so large that none of them individually is in a position to influence the price and output of the industry as a whole. The demand of individual buyer relative to the total demand is so small that he cannot influence the price of the product by his individual action.

Similarly, the supply of an individual seller is so small a fraction of the total output that he cannot influence the price of the product by his action alone. In other words, the individual seller is unable to influence the price of the product by increasing or decreasing its supply.

Rather, he adjusts his supply to the price of the product. He is “output adjuster”. Thus no buyer or seller can alter the price by his individual action. He has to accept the price for the product as fixed for the whole industry. He is a “price taker”.

Freedom of Entry or Exit of Firms

The next condition is that the firms should be free to enter or leave the industry. It implies that whenever the industry is earning excess profits, attracted by these profits some new firms enter the industry. In case of loss being sustained by the industry, some firms leave it.

Homogeneous Product

Each firm produces and sells a homogeneous product so that no buyer has any preference for the product of any individual seller over others. This is only possible if units of the same product produced by different sellers are perfect substitutes. In other words, the cross elasticity of the products of sellers is infinite.

No seller has an independent price policy. Commodities like salt, wheat, cotton and coal are homogeneous in nature. He cannot raise the price of his product. If he does so, his customers would leave him and buy the product from other sellers at the ruling lower price.

The above two conditions between themselves make the average revenue curve of the individual seller or firm perfectly

elastic, horizontal to the X-axis. It means that a firm can sell more or less at the ruling market price but cannot influence the price as the product is homogeneous and the number of sellers very large.

Absence of Artificial Restrictions

The next condition is that there is complete openness in buying and selling of goods. Sellers are free to sell their goods to any buyers and the buyers are free to buy from any sellers. In other words, there is no discrimination on the part of buyers or sellers.

Moreover, prices are liable to change freely in response to demand-supply conditions. There are no efforts on the part of the producers, the government and other agencies to control the supply, demand or price of the products. The movement of prices is unfettered.

Profit Maximisation Goal

Every firm has only one goal of maximising its profits.

Perfect Mobility of Goods and Factors

Another requirement of perfect competition is the perfect mobility of goods and factors between industries. Goods are free to move to those places where they can fetch the highest price. Factors can also move from a low-paid to a high-paid industry.

Perfect Knowledge of Market Conditions

This condition implies a close contact between buyers and sellers. Buyers and sellers possess complete knowledge about the prices at which goods are being bought and sold, and of the prices at which others are prepared to buy and sell. They have also perfect knowledge of the place where the transactions are being carried on. Such perfect knowledge of market conditions forces the sellers to sell their product at the prevailing market price and the buyers to buy at that price.

Perfect Competition vs Pure Competition

Perfect competition is often distinguished from pure competition, but they differ only in degree. The first five conditions

relate to pure competition while the remaining four conditions are also required for the existence of perfect competition. According to Chamberlin, pure competition means, competition unalloyed with monopoly elements," whereas perfect competition involves perfection in many other respects than in the absence of monopoly."

The practical importance of perfect competition is not much in the present times for few markets are perfectly competitive except those for staple food products and raw materials. That is why, Chamberlin says that perfect competition is a rare phenomenon." Though the real world does not fulfil the conditions of perfect competition, yet perfect competition is studied for the simple reason that it helps us in understanding the working of an economy, where competitive behaviour leads to the best allocation of resources and the most efficient organisation of production. A hypothetical model of a perfectly competitive industry provides the basis for appraising the actual working of economic institutions and organisations in any economy.

PRICE AND OUTPUT DETERMINATION

Price discrimination is a microeconomic pricing strategy where identical or largely similar goods or services are transacted at different prices by the same provider in different markets. Price discrimination is distinguished from product differentiation by the more substantial difference in production cost for the differently priced products involved in the latter strategy. Price differentiation essentially relies on the variation in the customers' willingness to pay and in the elasticity of their demand. Price discrimination, very differently, relies on monopoly power, including market share, product uniqueness, sole pricing power, etc.

The term differential pricing is also used to describe the practice of charging different prices to different buyers for the same quality and quantity of a product, but it can also refer to a combination of price differentiation and product differentiation. Other terms used to refer to price discrimination include equity pricing, preferential pricing, dual pricing and tiered pricing. Within the broader domain of price differentiation, a commonly accepted classification dating to the 1920s is:

- Personalized pricing (or first-degree price differentiation) – selling to each customer at a different price; this is also called one-to-one marketing. The optimal incarnation of this is called perfect price discrimination and maximizes the price that each customer is willing to pay.
- Product versioning or simply versioning (or second-degree price differentiation) – offering a product line by creating slightly different products for the purpose of price differentiation, i.e. a *vertical* product line. Another name given to versioning is menu pricing.
- Group pricing (or third-degree price differentiation) – dividing the market into segments and charging a different price to each segment (but the same price to each member of that segment). This is essentially a heuristic approximation that simplifies the problem in face of the difficulties with personalized pricing. Typical examples include student discounts and seniors' discounts.

Theoretical basis

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

The effects of price discrimination on social efficiency are unclear. Output can be expanded when price discrimination is very efficient. Even if output remains constant, price discrimination can reduce efficiency by misallocating output among consumers.

Price discrimination requires market segmentation and some means to discourage discount customers from becoming resellers

and, by extension, competitors. This usually entails using one or more means of preventing any resale: keeping the different price groups separate, making price comparisons difficult, or restricting pricing information. The boundary set up by the marketer to keep segments separate is referred to as a *rate fence*. Price discrimination is thus very common in services where resale is not possible; an example is student discounts at museums: In theory, students, for their condition as students, may get lower prices than the rest of the population for a certain product or service, and later will not become resellers, since what they received, may only be used or consumed by them. Another example of price discrimination is intellectual property, enforced by law and by technology. In the market for DVDs, laws require DVD players to be designed and produced with hardware or software that prevents inexpensive copying or playing of content purchased legally elsewhere in the world at a lower price. In the US the Digital Millennium Copyright Act has provisions to outlaw circumventing of such devices to protect the enhanced monopoly profits that copyright holders can obtain from price discrimination against higher price market segments.

Price discrimination can also be seen where the requirement that goods be identical is relaxed. For example, so-called “premium products” (including relatively simple products, such as cappuccino compared to regular coffee with cream) have a price differential that is not explained by the cost of production. Some economists have argued that this is a form of price discrimination exercised by providing a means for consumers to reveal their willingness to pay.

First degree

Exercising first degree (or perfect or primary) price discrimination requires the monopoly seller of a good or service to know the absolute maximum price (or reservation price) that every consumer is willing to pay. By knowing the reservation price, the seller is able to sell the good or service to each consumer at the maximum price they are willing to pay, and thus transform the consumer surplus into revenues, leading it to be the most

profitable form of price discrimination. So the profit is equal to the sum of consumer surplus and producer surplus. The marginal consumer is the one whose reservation price equals the marginal cost of the product. The seller produces more of their product than they would to achieve monopoly profits with no price discrimination, which means that there is no deadweight loss. Examples of this might be observed in markets where consumers bid for tenders, though, in this case, the practice of collusive tendering could reduce the market efficiency.

Second degree

In second-degree price discrimination, price varies according to quantity demanded. Larger quantities are available at a lower unit price. This is particularly widespread in sales to industrial customers, where bulk buyers enjoy discounts.

Additionally to second-degree price discrimination, sellers are not able to differentiate between different types of consumers. Thus, the suppliers will provide incentives for the consumers to differentiate themselves according to preference, which is done by quantity “discounts”, or non-linear pricing. This allows the supplier to set different prices to the different groups and capture a larger portion of the total market surplus.

In reality, different pricing may apply to differences in product quality as well as quantity. For example, airlines often offer multiple classes of seats on flights, such as first-class and economy class, with the first-class passengers receiving wine, beer and spirits with their ticket and the economy passengers offered only juice, pop, and water. This is a way to differentiate consumers based on preference, and therefore allows the airline to capture more consumer’s surplus.

Third degree

Third degree price discrimination means charging a different price to different consumer groups. For example, rail and tube (subway) travellers can be subdivided into commuter and casual travellers, and cinema goers can be subdivided into adults and children, with some theatres also offering discounts to full-time

students and seniors. Splitting the market into peak and off peak use of a service is very common and occurs with gas, electricity, and telephone supply, as well as gym membership and parking charges. Some parking lots charge less for “early bird” customers who arrive at the parking lot before a certain time.

(Some of these examples are not pure “price discrimination”, in that the differential price is related to production costs: the marginal cost of providing electricity or car parking spaces is very low outside peak hours. Incentivizing consumers to switch to off-peak usage is done as much to minimize costs as to maximize revenue.)

Modern taxonomy

The first/second/third degree taxonomy of price discrimination is due to Pigou (*Economics of Welfare*, 4th edition, 1932). The modern taxonomy of price discrimination. However, these categories are not mutually exclusive or exhaustive. Ivan Png (*Managerial Economics*, 2nd edition, 2002) suggests an alternative taxonomy:

Complete discrimination where each user purchases up to the point where the user’s marginal benefit equals the marginal cost of the item;

Direct segmentation where the seller can condition price on some attribute (like age or gender) that *directly* segments the buyers;

Indirect segmentation where the seller relies on some proxy (e.g., package size, usage quantity, coupon) to structure a choice that *indirectly* segments the buyers.

The hierarchy—complete/direct/indirect—is in decreasing order of profitability and information requirement. Complete price discrimination is most profitable, and requires the seller to have the most information about buyers. Indirect segmentation is least profitable, and requires the seller to have the least information about buyers.

Two part tariff

The two-part tariff is another form of price discrimination where the producer charges an initial fee then a secondary fee for

the use of the product. An example of this is razors, you pay an initial cost for the razor and then pay for the replacement blades. This pricing strategy works because it shifts the demand curve to the right: since you have already paid for the initial blade holder you will buy the blades which are now cheaper than buying a disposable razor.

Explanation

The purpose of price discrimination is generally to capture the market's consumer surplus. This surplus arises because, in a market with a single clearing price, some customers (the very low price elasticity segment) would have been prepared to pay more than the single market price. Price discrimination transfers some of this surplus from the consumer to the producer/marketer. Strictly, a consumer surplus need not exist, for example where some below-cost selling is beneficial due to fixed costs or economies of scale. An example is a high-speed internet connection shared by two consumers in a single building; if one is willing to pay less than half the cost, and the other willing to make up the rest but not to pay the entire cost, then price discrimination is necessary for the purchase to take place.

It can be proved mathematically that a firm facing a downward sloping demand curve that is convex to the origin will always obtain higher revenues under price discrimination than under a single price strategy. This can also be shown geometrically.

A single price (P) is available to all customers. The amount of revenue is represented by area P, A, Q, O . The consumer surplus is the area above line segment P, A but below the demand curve (D).

With price discrimination, the demand curve is divided into two segments (D_1 and D_2). A higher price (P_1) is charged to the low elasticity segment, and a lower price (P_2) is charged to the high elasticity segment. The total revenue from the first segment is equal to the area P_1, B, Q_1, O . The total revenue from the second segment is equal to the area E, C, Q_2, Q_1 . The sum of these areas will always be greater than the area without discrimination assuming the demand curve resembles a rectangular hyperbola

with unitary elasticity. The more prices that are introduced, the greater the sum of the revenue areas, and the more of the consumer surplus is captured by the producer.

Note that the above requires both first and second degree price discrimination: the right segment corresponds partly to different people than the left segment, partly to the same people, willing to buy more if the product is cheaper.

It is very useful for the price discriminator to determine the optimum prices in each market segment. Each segment is considered as a separate market with its own demand curve. As usual, the profit maximizing output (Q_t) is determined by the intersection of the marginal cost curve (MC) with the marginal revenue curve for the total market (MRT).

The firm decides what amount of the total output to sell in each market by looking at the intersection of marginal cost with marginal revenue (profit maximization). This output is then divided between the two markets, at the equilibrium marginal revenue level. Therefore, the optimum outputs are Q_a and Q_b . From the demand curve in each market we can determine the profit maximizing prices of P_a and P_b .

It is also important to note that the marginal revenue in both markets at the optimal output levels must be equal, otherwise the firm could profit from transferring output over to whichever market is offering higher marginal revenue.

Given that Market 1 has a price elasticity of demand of E_1 and Market 2 of E_2 , the optimal pricing ratio in Market 1 versus Market 2 is $\{ \displaystyle P_1/P_2 = [1+1/E_2]/[1+1/E_1] \}$.

IMPORTANCE OF PERFECT COMPETITION

Idealizing conditions of perfect competition

There is a set of market conditions which are assumed to prevail in the discussion of what perfect competition might be if it were theoretically possible to ever obtain such perfect market conditions. These conditions include:

- A large number of buyers and sellers – A large number of consumers with the willingness and ability to buy the

product at a certain price, and a large number of producers with the willingness and ability to supply the product at a certain price.

- Perfect information – All consumers and producers know all prices of products and utilities they would get from owning each product.
- Homogeneous products – The products are perfect substitutes for each other, (i.e., the qualities and characteristics of a market good or service do not vary between different suppliers).
- Well defined property rights – These determine what may be sold, as well as what rights are conferred on the buyer.
- No barriers to entry or exit
- Every participant is a price taker – No participant with market power to set prices
- Perfect factor mobility – In the long run factors of production are perfectly mobile, allowing free long term adjustments to changing market conditions.
- Profit maximization of sellers – Firms sell where the most profit is generated, where marginal costs meet marginal revenue.
- Rational buyers: Buyers make all trades that increase their economic utility and make no trades that do not increase their utility.
- No externalities – Costs or benefits of an activity do not affect third parties. This criteria also excludes any government intervention.
- Zero transaction costs – Buyers and sellers do not incur costs in making an exchange of goods in a perfectly competitive market.
- Non-increasing returns to scale and no network effects – The lack of economies of scale or network effects ensures that there will always be a sufficient number of firms in the industry.
- Anti-competitive regulation - It is assumed that a market of perfect competition shall provide the regulations and

protections implicit in the control of and elimination of anti-competitive activity in the market place.

Normal profit

In a perfect market the sellers operate at zero economic surplus: sellers make a level of return on investment known as normal profits.

Normal profit is a component of (implicit) costs and not a component of business profit at all. It represents the opportunity cost, as the time that the owner spends running the firm could be spent on running a different firm. The enterprise component of normal profit is thus the profit that a business owner considers necessary to make running the business worth her or his while i.e. it is comparable to the next best amount the entrepreneur could earn doing another job. Particularly if enterprise is not included as a factor of production, it can also be viewed a return to capital for investors including the entrepreneur, equivalent to the return the capital owner could have expected (in a safe investment), plus compensation for risk. In other words, the cost of normal profit varies both within and across industries; it is commensurate with the riskiness associated with each type of investment, as per the risk-return spectrum.

Only normal profits arise in circumstances of perfect competition when long run economic equilibrium is reached; there is no incentive for firms to either enter or leave the industry.

In competitive and contestable markets

Economic profit does not occur in perfect competition in long run equilibrium; if it did, there would be an incentive for new firms to enter the industry, aided by a lack of barriers to entry until there was no longer any economic profit. As new firms enter the industry, they increase the supply of the product available in the market, and these new firms are forced to charge a lower price to entice consumers to buy the additional supply these new firms are supplying as the firms all compete for customers. Incumbent firms within the industry face losing their existing customers to the new firms entering the industry, and are therefore forced to

lower their prices to match the lower prices set by the new firms. New firms will continue to enter the industry until the price of the product is lowered to the point that it is the same as the average cost of producing the product, and all of the economic profit disappears. When this happens, economic agents outside of the industry find no advantage to forming new firms that enter into the industry, the supply of the product stops increasing, and the price charged for the product stabilizes, settling into an equilibrium.

The same is likewise true of the long run equilibria of monopolistically competitive industries and, more generally, any market which is held to be contestable. Normally, a firm that introduces a differentiated product can initially secure a *temporary* market power for a *short while*. At this stage, the initial price the consumer must pay for the product is high, and the demand for, as well as the availability of the product in the market, will be limited. In the long run, however, when the profitability of the product is well established, and because there are few barriers to entry, the number of firms that produce this product will increase until the available supply of the product eventually becomes relatively large, the price of the product shrinks down to the level of the average cost of producing the product. When this finally occurs, all monopoly profit associated with producing and selling the product disappears, and the initial monopoly turns into a competitive industry. In the case of contestable markets, the cycle is often ended with the departure of the former “hit and run” entrants to the market, returning the industry to its previous state, just with a lower price and no economic profit for the incumbent firms. Profit can, however, occur in competitive and contestable markets in the short run, as firms jostle for market position. Once risk is accounted for, long-lasting economic profit in a competitive market is thus viewed as the result of constant cost-cutting and performance improvement ahead of industry competitors, allowing costs to be below the market-set price.

In uncompetitive markets

Economic profit is, however, much more prevalent in uncompetitive markets such as in a perfect monopoly or oligopoly

situation. In these scenarios, individual firms have some element of market power: Though monopolists are constrained by consumer demand, they are not price takers, but instead either price-setters or quantity setters. This allows the firm to set a price which is higher than that which would be found in a similar but more competitive industry, allowing them economic profit in both the long and short run.

The existence of economic profits depends on the prevalence of barriers to entry: these stop other firms from entering into the industry and sapping away profits, like they would in a more competitive market. In cases where barriers are present, but more than one firm, firms can collude to limit production, thereby restricting supply in order to ensure the price of the product remains high enough to ensure all of the firms in the industry achieve an economic profit.

However, some economists, for instance Steve Keen, a professor at the University of Western Sydney, argue that even an infinitesimal amount of market power can allow a firm to produce a profit and that the absence of economic profit in an industry, or even merely that some production occurs at a loss, in and of itself constitutes a barrier to entry.

In a single-goods case, a positive economic profit happens when the firm's average cost is less than the price of the product or service at the profit-maximizing output. The economic profit is equal to the quantity of output multiplied by the difference between the average cost and the price.

Government intervention

Often, governments will try to intervene in uncompetitive markets to make them more competitive. Antitrust (US) or competition (elsewhere) laws were created to prevent powerful firms from using their economic power to artificially create the barriers to entry they need to protect their economic profits. This includes the use of predatory pricing toward smaller competitors. For example, in the United States, Microsoft Corporation was initially convicted of breaking Anti-Trust Law and engaging in anti-competitive behavior in order to form one such barrier in

United States v. Microsoft; after a successful appeal on technical grounds, Microsoft agreed to a settlement with the Department of Justice in which they were faced with stringent oversight procedures and explicit requirements designed to prevent this predatory behaviour. With lower barriers, new firms can enter the market again, making the long run equilibrium much more like that of a competitive industry, with no economic profit for firms.

If a government feels it is impractical to have a competitive market – such as in the case of a natural monopoly – it will sometimes try to regulate the existing uncompetitive market by controlling the price firms charge for their product. For example, the old AT&T (regulated) monopoly, which existed before the courts ordered its breakup, had to get government approval to raise its prices. The government examined the monopoly's costs, and determined whether or not the monopoly should be able raise its price and if the government felt that the cost did not justify a higher price, it rejected the monopoly's application for a higher price. Although a regulated firm will not have an economic profit as large as it would in an unregulated situation, it can still make profits well above a competitive firm in a truly competitive market.

Results

In a perfectly competitive market, the demand curve facing a firm is perfectly elastic.

As mentioned above, the perfect competition model, if interpreted as applying also to short-period or very-short-period behaviour, is approximated only by markets of homogeneous products produced and purchased by very many sellers and buyers, usually organized markets for agricultural products or raw materials. In real-world markets, assumptions such as perfect information cannot be verified and are only approximated in organized double-auction markets where most agents wait and observe the behaviour of prices before deciding to exchange (but in the long-period interpretation perfect information is not necessary, the analysis only aims at determining the average around which market prices gravitate, and for gravitation to operate one does not need perfect information).

In the absence of externalities and public goods, perfectly competitive equilibria are Pareto-efficient, i.e. no improvement in the utility of a consumer is possible without a worsening of the utility of some other consumer. This is called the First Theorem of Welfare Economics. The basic reason is that no productive factor with a non-zero marginal product is left unutilized, and the units of each factor are so allocated as to yield the same indirect marginal utility in all uses, a basic efficiency condition (if this indirect marginal utility were higher in one use than in other ones, a Pareto improvement could be achieved by transferring a small amount of the factor to the use where it yields a higher marginal utility).

A simple proof assuming differentiable utility functions and production functions is the following. Let w_j be the 'price' (the rental) of a certain factor j , let MP_{j1} and MP_{j2} be its marginal product in the production of goods 1 and 2, and let p_1 and p_2 be these goods' prices. In equilibrium these prices must equal the respective marginal costs MC_1 and MC_2 ; remember that marginal cost equals factor 'price' divided by factor marginal productivity (because increasing the production of good by one very small unit through an increase of the employment of factor j requires increasing the factor employment by $1/MP_{ji}$ and thus increasing the cost by w_j/MP_{ji} , and through the condition of cost minimization that marginal products must be proportional to factor 'prices' it can be shown that the cost increase is the same if the output increase is obtained by optimally varying all factors). Optimal factor employment by a price-taking firm requires equality of factor rental and factor marginal revenue product, $w_j = p_i MP_{ji}$, so we obtain $p_1 = MC_1 = w_j / MP_{j1}$, $p_2 = MC_2 = w_j / MP_{j2}$.

Now choose any consumer purchasing both goods, and measure his utility in such units that in equilibrium his marginal utility of money (the increase in utility due to the last unit of money spent on each good), $MU_1/p_1 = MU_2/p_2$, is 1. Then $p_1 = MU_1$, $p_2 = MU_2$. The indirect marginal utility of the factor is the increase in the utility of our consumer achieved by an increase in the employment of the factor by one (very small) unit; this increase in utility through allocating the small increase in factor utilization

to good 1 is $MP_{j1}MU_1=MP_{j1}p_1=w_j$, and through allocating it to good 2 it is $MP_{j2}MU_2=MP_{j2}p_2=w_j$, again. With our choice of units the marginal utility of the amount of the factor consumed directly by the optimizing consumer is again w , so the amount supplied of the factor too satisfies the condition of optimal allocation.

Monopoly violates this optimal allocation condition, because in a monopolized industry market price is above marginal cost, and this means that factors are underutilized in the monopolized industry, they have a higher indirect marginal utility than in their uses in competitive industries. Of course this theorem is considered irrelevant by economists who do not believe that general equilibrium theory correctly predicts the functioning of market economies; but it is given great importance by neoclassical economists and it is the theoretical reason given by them for combating monopolies and for antitrust legislation.

Shutdown point

In the short run, a firm operating at a loss [$R < TC$ (revenue less than total cost) or $P < ATC$ (price less than unit cost)] must decide whether to continue to operate or temporarily shut down. The shutdown rule states “in the short run a firm should continue to operate if price exceeds average variable costs”. Restated, the rule is that for a firm to continue producing in the short run it must earn sufficient revenue to cover its variable costs. The rationale for the rule is straightforward: By shutting down a firm avoids all variable costs. However, the firm must still pay fixed costs. Because fixed costs must be paid regardless of whether a firm operates they should not be considered in deciding whether to produce or shut down. Thus in determining whether to shut down a firm should compare total revenue to total variable costs (VC) rather than total costs (FC + VC). If the revenue the firm is receiving is greater than its total variable cost ($R > VC$), then the firm is covering all variable costs and there is additional revenue (“contribution”), which can be applied to fixed costs. (The size of the fixed costs is irrelevant as it is a sunk cost. The same consideration is used whether fixed costs are one dollar or one million dollars.) On the other hand, if $VC > R$ then the firm is not

covering its production costs and it should immediately shut down. The rule is conventionally stated in terms of price (average revenue) and average variable costs. The rules are equivalent (if one divides both sides of inequality $TR > TVC$ by Q gives $P > AVC$). If the firm decides to operate, the firm will continue to produce where marginal revenue equals marginal costs because these conditions insure not only profit maximization (loss minimization) but also maximum contribution.

Another way to state the rule is that a firm should compare the profits from operating to those realized if it shut down and select the option that produces the greater profit. A firm that is shut down is generating zero revenue and incurring no variable costs. However, the firm still has to pay fixed cost. So the firm's profit equals fixed costs or "FC. An operating firm is generating revenue, incurring variable costs and paying fixed costs. The operating firm's profit is $R - VC - FC$. The firm should continue to operate if $R - VC - FC \geq -FC$, which simplified is $R \geq VC$. The difference between revenue, R , and variable costs, VC , is the contribution to fixed costs and any contribution is better than none. Thus, if $R \geq VC$ then firm should operate. If $R < VC$ the firm should shut down.

A decision to shut down means that the firm is temporarily suspending production. It does not mean that the firm is going out of business (exiting the industry). If market conditions improve, and prices increase, the firm can resume production. Shutting down is a short-run decision. A firm that has shut down is not producing. The firm still retains its capital assets; however, the firm cannot leave the industry or avoid its fixed costs in the short run. Exit is a long-term decision. A firm that has exited an industry has avoided all commitments and freed all capital for use in more profitable enterprises.

However, a firm cannot continue to incur losses indefinitely. In the long run, the firm will have to earn sufficient revenue to cover all its expenses and must decide whether to continue in business or to leave the industry and pursue profits elsewhere. The long-run decision is based on the relationship of the price and long-run average costs. If $P \geq AC$ then the firm will not exit the

industry. If $P < AC$, then the firm will exit the industry. These comparisons will be made after the firm has made the necessary and feasible long-term adjustments. In the long run a firm operates where marginal revenue equals long-run marginal costs.

Short-run supply curve

The short-run (SR) supply curve for a perfectly competitive firm is the marginal cost (MC) curve at and above the shutdown point. Portions of the marginal cost curve below the shutdown point are not part of the SR supply curve because the firm is not producing any positive quantity in that range. Technically the SR supply curve is a discontinuous function composed of the segment of the MC curve at and above minimum of the average variable cost curve and a segment that runs on the vertical axis from the origin to but not including a point at the height of the minimum average variable cost.

Examples

Though there is no actual perfectly competitive market in the real world, a number of approximations exist:

An example is that of a large action of identical goods with all potential buyers and sellers present. By design, a stock exchange resembles this, not as a complete description (for no markets may satisfy all requirements of the model) but as an approximation. The flaw in considering the stock exchange as an example of Perfect Competition is the fact that large institutional investors (e.g. investment banks) may solely influence the market price. This, of course, violates the condition that “no one seller can influence market price”.

Horse betting is also quite a close approximation. When placing bets, consumers can just look down the line to see who is offering the best odds, and so no one bookie can offer worse odds than those being offered by the market as a whole, since consumers will just go to another bookie. This makes the bookies price-takers. Furthermore, the product on offer is very homogeneous, with the only differences between individual bets being the pay-off and the horse. Of course, there are not an infinite amount of bookies, and

some barriers to entry exist, such as a license and the capital required to set up.

Criticisms

The use of the assumption of perfect competition as the foundation of price theory for product markets is often criticized as representing all agents as passive, thus removing the active attempts to increase one's welfare or profits by price undercutting, product design, advertising, innovation, activities that – the critics argue – characterize most industries and markets. These criticisms point to the frequent lack of realism of the assumptions of product homogeneity and impossibility to differentiate it, but apart from this the accusation of passivity appears correct only for short-period or very-short-period analyses, in long-period analyses the inability of price to diverge from the natural or long-period price is due to active reactions of entry or exit.

Some economists have a different kind of criticism concerning perfect competition model. They are not criticizing the price taker assumption because it makes economic agents too “passive”, but because it then raises the question of who sets the prices. Indeed, if everyone is price taker, there is the need for a benevolent planner who gives and sets the prices, in other word, there is a need for a “price maker”. Therefore, it makes the perfect competition model appropriate not to describe a decentralize “market” economy but a centralized one. This in turn means that such kind of model has more to do with communism than capitalism.

Another frequent criticism is that it is often not true that in the short run differences between supply and demand cause changes in price; especially in manufacturing, the more common behaviour is alteration of production without nearly any alteration of price.

The critics of the assumption of perfect competition in product markets seldom question the basic neoclassical view of the working of market economies for this reason. The Austrian School insists strongly on this criticism, and yet the neoclassical view of the working of market economies as fundamentally efficient, reflecting consumer choices and assigning to each agent his contribution to

social welfare, is esteemed to be fundamentally correct. Some non-neoclassical schools, like Post-Keynesians, reject the neoclassical approach to value and distribution, but not because of their rejection of perfect competition as a reasonable approximation to the working of most product markets; the reasons for rejection of the neoclassical 'vision' are different views of the determinants of income distribution and of aggregated demand.

In particular, the rejection of perfect competition does not generally entail the rejection of free competition as characterizing most product markets; indeed it has been argued that competition is stronger nowadays than in 19th century capitalism, owing to the increasing capacity of big conglomerate firms to enter any industry: therefore the classical idea of a tendency toward a uniform rate of return on investment in all industries owing to free entry is even more valid today; and the reason why General Motors, Exxon or Nestlé do not enter the computers or pharmaceutical industries is not insurmountable barriers to entry but rather that the rate of return in the latter industries is already sufficiently in line with the average rate of return elsewhere as not to justify entry. On this few economists, it would seem, would disagree, even among the neoclassical ones. Thus when the issue is normal, or long-period, product prices, differences on the validity of the perfect competition assumption do not appear to imply important differences on the existence or not of a tendency of rates of return toward uniformity as long as entry is possible, and what is found fundamentally lacking in the perfect competition model is the absence of marketing expenses and innovation as causes of costs that do enter normal average cost.

The issue is different with respect to factor markets. Here the acceptance or denial of perfect competition in labour markets does make a big difference to the view of the working of market economies. One must distinguish neoclassical from non-neoclassical economists. For the former, absence of perfect competition in labour markets, e.g. due to the existence of trade unions, impedes the smooth working of competition, which if left free to operate would cause a decrease of wages as long as there were unemployment, and would finally ensure the full employment of

labour: labour unemployment is due to absence of perfect competition in labour markets. Most non-neoclassical economists deny that a full flexibility of wages would ensure the full employment of labour and find a stickiness of wages an indispensable component of a market economy, without which the economy would lack the regularity and persistence indispensable to its smooth working. This was, for example, John Maynard Keynes's opinion.

Particularly radical is the view of the Sraffian school on this issue: the labour demand curve cannot be determined hence a level of wages ensuring the equality between supply and demand for labour does not exist, and economics should resume the viewpoint of the classical economists, according to whom competition in labour markets does not and cannot mean indefinite price flexibility as long as supply and demand are unequal, it only means a tendency to equality of wages for similar work, but the level of wages is necessarily determined by complex sociopolitical elements; custom, feelings of justice, informal allegiances to classes, as well as overt coalitions such as trade unions, far from being impediments to a smooth working of labour markets that would be able to determine wages even without these elements, are on the contrary indispensable because without them there would be no way to determine wages.

MONOPOLY: DISCRIMINATION MONOPOLY, DUMPING, COMPARISON BETWEEN PERFECT COMPETITION AND MONOPOLY

Characteristics of Monopoly

The main features of monopoly are as follows:

1. Under monopoly, there is one producer or seller of a particular product and there is no difference between a firm and an industry. Under monopoly a firm itself is an industry.
2. A monopoly may be individual proprietorship or partnership or joint stock company or a co-operative society or a government company.

3. A monopolist has full control on the supply of a product. Hence, the elasticity of demand for a monopolist's product is zero.
4. There is no close substitute of a monopolist's product in the market. Hence, under monopoly, the cross elasticity of demand for a monopoly product with some other good is very low.
5. There are restrictions on the entry of other firms in the area of monopoly product.
6. A monopolist can influence the price of a product. He is a price-maker, not a price-taker.
7. Pure monopoly is not found in the real world.
8. Monopolist cannot determine both the price and quantity of a product simultaneously.
9. Monopolist's demand curve slopes downwards to the right. That is why, a monopolist can increase his sales only by decreasing the price of his product and thereby maximise his profit. The marginal revenue curve of a monopolist is below the average revenue curve and it falls faster than the average revenue curve. This is because a monopolist has to cut down the price of his product to sell an additional unit.

Price Discrimination In Monopoly

Price discrimination may be (a) personal, (b) local, or (c) according to trade or use:

- (a) *Personal*: It is personal when different prices are charged for different persons.
- (b) *Local*: It is local when the price varies according to locality.
- (c) *According to Trade or Use*: It is according to trade or use when different prices are charged for different uses to which the commodity is put, for example, electricity is supplied at cheaper rates for domestic than for commercial purposes.

Some monopolists used product differentiation for price discrimination by means of special labels, wrappers, packing, etc.

For example, the perfume manufacturers discriminate prices of the same fragrance by packing it with different labels or brands.

Conditions of Price-Discrimination: There are three main types of situation:

- (a) *When consumers have certain preferences or prejudices.* Certain consumers usually have the irrational feeling that they are paying higher prices for a good because it is of a better quality, although actually it may be of the same quality. Sometimes, the price differences may be so small that consumers do not consider it worthwhile to bother about such differences.
- (b) *When the nature of the good is such as makes it possible for the monopolist to charge different prices.* This happens particularly when the good in question is a direct service.
- (c) *When consumers are separated by distance or tariff barriers.* A good may be sold in one town for Re. 1 and in another town for Rs. 2. Similarly, the monopolist can charge higher prices in a city with greater distance or a country levying heavy import duty.

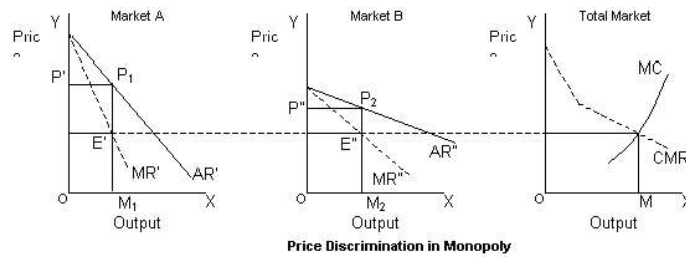
Conditions making Price Discrimination Possible and Profitable: The following conditions are essential to make price discrimination possible and profitable:

- (a) *The elasticities of demand in different markets must be different.* The market is divided into sub-markets. The sub-market will be arranged in ascending order of their elasticities, the higher price being charged in the least elastic market and vice versa.
- (b) *The costs incurred in dividing the market into sub-markets and keeping them separate should not be so large as to neutralise the difference in demand elasticities.*
- (c) *There should be complete agreement among the sellers otherwise the competitors will gain by selling in the dear market.*
- (d) *When goods are sold on special orders* because then the purchaser cannot know what is being charged from others.

Price Determination under Price Discrimination:

First of all, the monopolist divides his total market into sub-

markets. In the following diagrams, the monopolist has divided his total market into two sub-markets, i.e., A and B:



The monopolist has now to decide at what level of output he should produce. To achieve maximum profit, hence, he will be in equilibrium at output at which $MR=MC$, and MC curve cuts the MR curve from below.

The equilibrium of the discriminating monopolist is established at output OM at which MC cuts CMR . The output OM is distributed between two markets in such a way that marginal revenue in each is equal to ME .

Therefore, he will sell output OM_1 in Market A, because only at this output marginal revenue MR' in Market A is equal to ME ($M_1E' = ME$). The same condition is applied in Market B where MR'' is equal to ME ($M_2E'' = ME$). It is also shown that in Market B in which elasticity of demand is greater, the price charged is lower than that in Market A where the elasticity of demand is less.

Price and Output Determination under Monopoly

Monopoly refers to a market structure in which there is a single producer or seller that has a control on the entire market. This single seller deals in the products that have no close substitutes and has a direct demand, supply, and prices of a product. Therefore, in monopoly, there is no distinction between an one organization constitutes the whole industry.

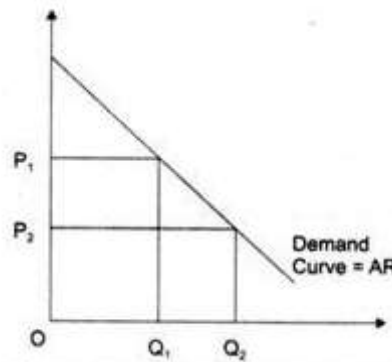
Demand and Revenue under Monopoly

In monopoly, there is only one producer of a product, who influences the price of the product by making Change in supply.

The producer under monopoly is called monopolist. If the monopolist wants to sell more, he/she can reduce the price of a product. On the other hand, if he/she is willing to sell less, he/she can increase the price.

As we know, there is no difference between organization and industry under monopoly. Accordingly, the demand curve of the organization constitutes the demand curve of the entire industry. The demand curve of the monopolist is Average Revenue (AR), which slopes downward.

AR curve of the monopolist:



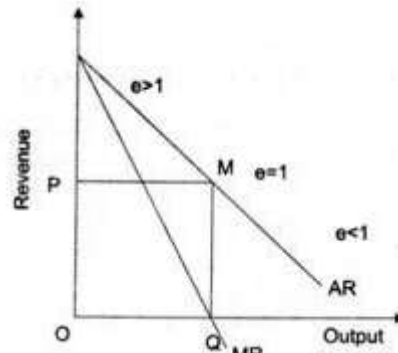
It can be seen that more quantity (OQ_2) can only be sold at lower price (OP_2). Under monopoly, the slope of AR curve is downward, which implies that if the high prices are set by the monopolist, the demand will fall. In addition, in monopoly, AR curve and Marginal Revenue (MR) curve are different from each other. However, both of them slope downward.

The negative AR and MR curve depicts the following facts:

- i. When MR is greater than AR, the AR rises
- ii. When MR is equal to AR, then AR remains constant
- iii. When MR is lesser than AR, then AR falls

Here, AR is the price of a product, As we know, AR falls under monopoly; thus, MR is less than AR.

AR and MR curves under monopoly:



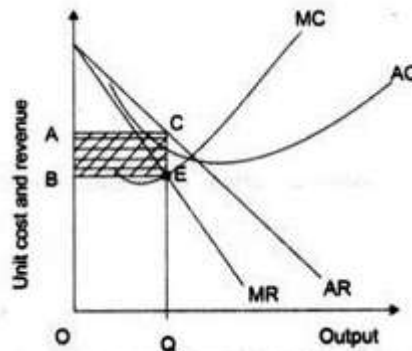
MR curve is shown below the AR curve because AR falls.

Monopoly Equilibrium

Single organization constitutes the whole industry in monopoly. Thus, there is no need for separate analysis of equilibrium of organization and industry in case of monopoly. The main aim of monopolist is to earn maximum profit as of a producer in perfect competition.

Unlike perfect competition, the equilibrium, under monopoly, is attained at the point where profit is maximum that is where $MR=MC$. Therefore, the monopolist will go on producing additional units of output as long as MR is greater than MC, to earn maximum profit.

Let us learn monopoly equilibrium:



If output is increased beyond OQ, MR will be less than MC.

Thus, if additional units are produced, the organization will incur loss. At equilibrium point, total profits earned are equal to shaded area ABEC. E is the equilibrium point at which $MR=MC$ with quantity as OQ.

It should be noted that under monopoly, price forms the following relation with the MC:

$$\text{Price} = AR$$

$$MR = AR [(e-1)/e]$$

e = Price elasticity of demand

As in equilibrium $MR=MC$

$$MC = AR [(e-1)/e]$$

Determining Price and Output under Monopoly:

Suppose demand function for monopoly is $Q = 200 - 0.4Q$

Price function is $P = 1000 - 10Q$

Cost function is $TC = 100 + 40Q + Q^2$

Maximum profit is achieved where $MR=MC$

To find MR, TR is derived.

$$TR = (1000 - 10Q) Q = 1000Q - 10Q^2$$

$$MR = \frac{TR}{Q} = 1000 - 20Q$$

$$MC = \frac{TC}{Q} = 40 + 2Q$$

$$MR = MC$$

$$1000 - 20Q = 40 + 2Q$$

$$Q = 43.63 \text{ (44 approx.)} = \text{Profit Maximizing Output}$$

$$\text{Profit maximizing price} = 1000 - 20 \times 44 = 120$$

$$\text{Total maximum profit} = TR - TC = (1000Q - 10Q^2) - (100 + 40Q + Q^2)$$

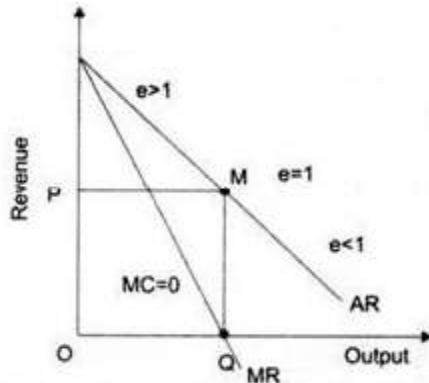
$$\text{At } Q = 44$$

$$\text{Total maximum profit} = \text{Rs. } 20844$$

Monopoly Equilibrium in Case of Zero Marginal Cost:

In certain situations, it may happen that MC is zero, which implies that the cost of production is zero. For example, cost of production of spring water is zero. However, the monopolist will

set its price to earn profit. The monopoly equilibrium when MC is zero:



AR is the average revenue curve and MR is the marginal revenue curve. In such a case, the total cost is zero; therefore, AR and MR are also zero. The equilibrium position is achieved at the point where MR equals zero that is at output OQ and price P. We can see that point M is the mid-point of AR curve, where elasticity of demand is unity. Therefore, when $MC = 0$, the equilibrium of the monopolist is established at the output (OQ) where elasticity of demand is unity.

Short-Run and Long-Run View under Monopoly

Till now, we have discussed monopoly equilibrium without taking into consideration the short-run and long-run period. This is because there is not so much difference under short run and long run analysis in monopoly.

In the short run, the monopolist should make sure that the price should not go below Average Variable Cost (AVC). The equilibrium under monopoly in long-run is same as in short-run. However, in long-run, the monopolist can expand the size of its plants according to demand. The adjustment is done to make MR equal to the long run MC.

In the long-run, under perfect competition, the equilibrium position is attained by entry or exit of the organizations. In monopoly, the entry of new organizations is restricted.

The monopolist may hold some patents or copyright that limits the entry of other players in the market. When a monopolist incurs losses, he/she may exit the business. On the other hand, if profits are earned, then he/she may increase the plant size to gain more profit.

MONOPOLISTIC COMPETITION

Monopolistic competition refers to a market situation where there are many firms selling a differentiated product. "There is competition which is keen, though not perfect, among many firms making very similar products." No firm can have any perceptible influence on the price-output policies of the other sellers nor can it be influenced much by their actions. Thus monopolistic competition refers to competition among a large number of sellers producing close but not perfect substitutes for each other.

Its Features:

The following are the main features of monopolistic competition:

- (1) Large Number of Sellers: In monopolistic competition the number of sellers is large. They are "many and small enough" but none controls a major portion of the total output. No seller by changing its price-output policy can have any perceptible effect on the sales of others and in turn be influenced by them. Thus there is no recognised interdependence of the price-output policies of the sellers and each seller pursues an independent course of action.
- (2) Product Differentiation: One of the most important features of the monopolistic competition is differentiation. Product differentiation implies that products are different in some ways from each other. They are heterogeneous rather than homogeneous so that each firm has an absolute monopoly in the production and sale of a differentiated product. There is, however, slight difference between one product and other in the same category.

Products are close substitutes with a high cross-elasticity and not perfect substitutes. Product "differentiation may

be based upon certain characteristics of the products itself, such as exclusive patented features; trade-marks; trade names; peculiarities of package or container, if any; or singularity in quality, design, colour, or style. It may also exist with respect to the conditions surrounding its sales."

- (3) Freedom of Entry and Exit of Firms: Another feature of monopolistic competition is the freedom of entry and exit of firms. As firms are of small size and are capable of producing close substitutes, they can leave or enter the industry or group in the long run.
- (4) Nature of Demand Curve: Under monopolistic competition no single firm controls more than a small portion of the total output of a product. No doubt there is an element of differentiation nevertheless the products are close substitutes. As a result, a reduction in its price will increase the sales of the firm but it will have little effect on the price-output conditions of other firms, each will lose only a few of its customers.

Likewise, an increase in its price will reduce its demand substantially but each of its rivals will attract only a few of its customers. Therefore, the demand curve (average revenue curve) of a firm under monopolistic competition slopes downward to the right. It is elastic but not perfectly elastic within a relevant range of prices of which he can sell any amount.

- (5) Independent Behaviour: In monopolistic competition, every firm has independent policy. Since the number of sellers is large, none controls a major portion of the total output. No seller by changing its price-output policy can have any perceptible effect on the sales of others and in turn be influenced by them.
- (6) Product Groups: There is no any 'industry' under monopolistic competition but a 'group' of firms producing similar products. Each firm produces a distinct product and is itself an industry. Chamberlin lumps together firms producing very closely related products and calls them product groups, such as cars, cigarettes, etc.

- (7) **Selling Costs:** Under monopolistic competition where the product is differentiated, selling costs are essential to push up the sales. Besides, advertisement, it includes expenses on salesman, allowances to sellers for window displays, free service, free sampling, premium coupons and gifts, etc.
- (8) **Non-price Competition:** Under monopolistic competition, a firm increases sales and profits of his product without a cut in the price. The monopolistic competitor can change his product either by varying its quality, packing, etc. or by changing promotional programmes.

PRICE DETERMINATION

Price Discrimination

Elasticity Conditions for Price Discrimination

In a competitive market, price discrimination occurs when identical goods and services are sold at different prices by the same provider.

Price Discrimination

In a competitive market, price discrimination occurs when identical goods and services are sold at different prices by the same provider. In pure price discrimination, the seller will charge the buyer the absolute maximum price that he is willing to pay. Companies use price discrimination in order to make the most revenue possible from every customer. This allows the producer to capture more of the total surplus by selling to consumers at prices closer to their maximum willingness to pay.

Price discrimination: A producer that can charge price P_a to its customers with inelastic demand and P_b to those with elastic demand can extract more total profit than if it had charged just one price. An example of price discrimination would be the cost of movie tickets. Prices at one theater are different for children, adults, and seniors. The prices of each ticket can also vary based on the day and chosen show time. Ticket prices also vary depending on the portion of the country as well.

Industries use price discrimination as a way to increase revenue. It is possible for some industries to offer retailers different prices based solely on the volume of products purchased. Price discrimination can also be based on age, location, desire for the product, and customer wage.

Forms of Price Discrimination

There are a variety of ways in which industries legally use price discrimination. It is not important that pricing information be restricted, or that the price discriminated groups be unaware that others are being charged different prices:

- **Coupons:** coupons are used in retail as a way to distinguish customers by their reserve price. The assumption is that individuals who collect coupons are more sensitive to a higher price than those who don't. By offering coupons, a producer can charge a higher price to price-insensitive customers and provide a discount to price-sensitive individuals.
- **Premium pricing:** premium products are priced at a level that is well beyond their marginal cost. For example, a regular cup of coffee might be priced at \$1, while a premium coffee is \$2.50.
- **Discounts based on occupation:** many businesses offer reduced prices to active military members. This can increase sales to the target group and provide positive publicity for the business which leads to increased sales. Less publicized discounts are also offered to off duty service workers such as police.
- **Retail incentives:** retail incentives are used to increase market share or revenues. They include rebates, bulk and quantity pricing, seasonal discounts
- **Gender based discounts:** gender based discounts are offered in some countries including the United States. Examples include free drinks at bars for women on "Ladies Night," men often receive lower prices at the dry cleaners and hair salons than women because women clothes and hair generally take more time to work with. In contrast, men

usually have higher car insurance rates than women based on the likelihood of being in an accident based on their age.

- Financial aid: financial aid is offered to college students based on either the student and/or the parents economic situation.
- Haggling: haggling is a form of price negotiation that requires knowledge and confidence from the customer.

Industries that Use Price Discrimination

The airline industry uses price discrimination regularly when they sell travel tickets simultaneously to different market segments. Price discrimination is evident within individual airlines, but also in the industry as a whole. Tickets vary based on the location within the plane, the time and day of the flight, the time of year, and what city the aircraft is traveling to. Prices can vary greatly within an airline and also among airlines. Customers must search for the best priced ticket based on their needs. Airlines do offer other forms of price discrimination including discounts, vouchers, and member perks for individuals with membership cards.

The pharmaceutical industry experiences international price discrimination. Drug manufacturers charge more for drugs in wealthier countries than in poor ones. For example, the United States has the highest drug prices in the world. On average, Europeans pay 56% less than Americans do for the same prescription medications. However, in many countries with lower drug costs, the difference in price is absorbed into the taxes which results in lower average salaries when compared to those in the United States.

Academic textbooks are another industry known for price discrimination. Textbooks in the United States are more expensive than they are overseas. Because most of the textbooks are published in the United States, it is obvious that transportation costs do not raise the price of the books. In the United States price discrimination on textbooks is due to copyright protection laws. Also, in the United States textbooks are mandatory where as in other countries they are viewed as optional study aids.

Analysis of Price Discrimination

Price discrimination is present in commerce when sellers adjust the price on the same product in order to make the most revenue possible.

Price Discrimination Criteria

Within commerce there are specific criteria that must be met in order for price discrimination to occur:

- The firm must have market power.
- The firm must be able to recognize differences in demand.
- The firm must have the ability to prevent arbitration, or resale of the product.

Price Discrimination

Price discrimination exists within a market when the sales of identical goods or services are sold at different prices by the same provider. The goal of price discrimination is for the seller to make the most profit possible. Although the cost of producing the products is the same, the seller has the ability to increase the price based on location, consumer financial status, product demand, etc.

Sales Revenue: These graphs shows the difference in sales revenue with and without price discrimination. The intent of price discrimination is for the seller to make the most profit possible.

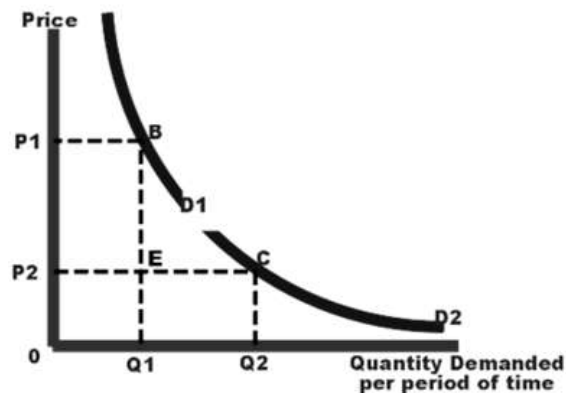
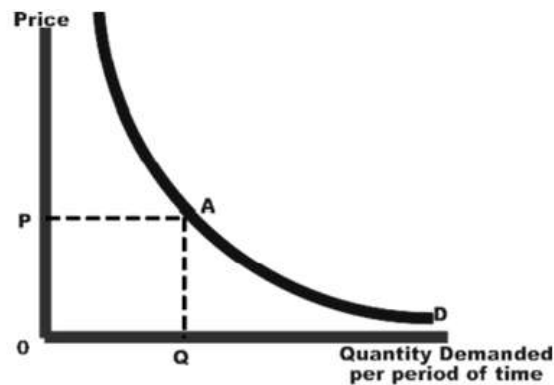
Types of Price Discrimination

In commerce there are three types of price discrimination that exist. The exact price discrimination method that is used depends on the factors within the particular market.

- First degree price discrimination: the monopoly seller of a good or service must know the absolute maximum price that every consumer is willing to pay and can charge each customer that exact amount. This allows the seller to obtain the highest revenue possible.
- Second degree price discrimination: the price of a good or service varies according to the quantity demanded. Larger

quantities are available at a lower price (higher discounts are given to consumers who buy a good in bulk quantities).

- Third degree price discrimination: the price varies according to consumer attributes such as age, sex, location, and economic status.



EXCESS CAPACITY

Theory of Excess Capacity

The doctrine of excess (or unutilised) capacity is associated with monopolistic competition in the long-run and is defined as "the difference between ideal (optimum) output and the output actually attained in the long-run."

Prof. Chamberlin's explanation of the theory of excess capacity is different from that of ideal (optimum) output under perfect competition. Under perfect competition, each firm produces at the minimum on its LAC curve and its horizontal demand curve is tangent to it at that point. Its output is ideal and there is no excess capacity in the long run.

Since under monopolistic competition the demand curve of the firm is downward sloping due to product differentiation, the long-run equilibrium of the firm is to the left of the minimum point on the LAC curve. This shows the existence of excess capacity.

According to Chamberlin, so long as there is freedom of entry and price competition in the product group under monopolistic competition, the tangency point between the firm's demand curve and the LAC curve would lead to the "ideal output". This is because consumers want product differentiation and they are willing to accept increased production costs in return for choice and variety of products that are available under monopolistic competition.

Chamberlin asserts that the difference between the actual long-run average cost of production of a monopolistic competitive firm with free entry and price competition and the minimum LAC represents the "cost of differentness" due to product differentiation. He does not regard this difference in average cost of production as a measure of "excess capacity".

Chamberlin's analysis of excess capacity can be divided into two parts: (1) Entry into the group with price competition; and (2) entry with non-price competition.

Assumptions

Chamberlin's concept of excess capacity assumes that:

- (i) The number of firms is large;
- (ii) Each produces a similar product independently of the others;
- (iii) It can charge a lower price and attract other's customers and by raising its price will lose some of its customers;

- (iv) Consumers preferences are fairly evenly distributed among the different varieties of products;
- (v) No firm has an institutional monopoly over the product;
- (vi) Firms are free to enter its field of production; and
- (vii) The long-run cost curves of all the firms are identical and are U-shaped.

Excess Capacity with Price Competition

Given these assumptions, Chamberlin's concepts of ideal output and excess capacity with active price competition. Suppose the initial short-run equilibrium is at point S where the demand curve of firms dd and the group demand curve DD intersect, and the existing firms are earning super-normal profits because the price OP corresponding to point S is above the LAC curve.

In the long-run, attracted by super-normal profits, new firms enter the group. They produce the similar product which reduces the sales of each firm in the group and pushes the group demand curve DD to $D'D'$. The new equilibrium is established at point A where the curve DD' is tangent to the LAC curve. Competition among the firms leads to price reductions and the $d'd'$ curve of each firm slides downwards to d_1d_1 along the $D'D'$ curve till it is tangent to the LAC curve at point A_1 .

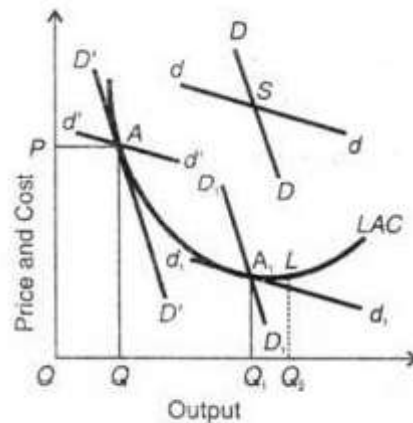


Fig. 6.

Simultaneously, the $D'D'$ curve is pushed down to D_1D_1 and it intersects both the d_1d_1 curve and the LAC curve at A_1 . This is the long-run stable equilibrium position of the group. Each firm is producing the ideal output OQ_1 at $Q_1 A_1$ price, earning normal profits and there is no excess capacity.

The ideal output under perfect competition is OQ_2 which is established at the minimum point L of the LAC curve. The output difference Q_1Q_2 between the perfectly competitive output and the monopolistic competitive output is the cost difference which consumers are willing to pay for enjoying a variety of products under product differentiation. Thus with free entry and active price competition under monopolistic competition, there is no excess capacity.

SELLING COSTS

Absence of Selling Costs

Under perfect competition, the costs of advertising, sales-promotion, etc. do not arise because all firms produce a homogeneous product.

Expenses of Selling Costs

Selling costs are the expenses on advertisement, salesmanship, free sampling, free service, door-to-door canvassing, and so on. There is no selling problem under perfect competition where the product is homogeneous.

The firm can sell at the ruling market price any quantity of its product. Therefore, there is no need for advertising. If, however, all firms want to sell more, competition among them will lead to price reduction until the new equilibrium price is reached. Each firm can sell as much as it wants to sell at this price.

Under monopoly also, selling costs are not required as there are no competitors. But the monopolist may sometimes advertise his product to acquaint the people about the use and price of his product so that they may continue to buy his product.

Under monopolistic competition where the product is differentiated, selling costs are essential to push up the sales. They

are incurred to persuade a buyer to purchase one product in preference to another. Chamberlin defines them "as costs incurred in order to alter the position or shape of the demand curve for a product.

He regards advertisement of all types as synonymous with selling costs. But in the present day business nomenclature, the term selling costs is wider than advertising and it includes besides advertising, expenses on salesmen, allowances to sellers for window displays, free service, free sampling, premium coupons and gifts, etc.

Advertisement is of two types: Informative and competitive.

Informative advertisement is meant to acquaint the buyers with the existence and uses of the product. Advertisements appearing in newspapers are of this type. They only give some general or technical information about the product without attempting to persuade buyers to buy their product.

This type of advertising is meant for the ordinary buyer to enable him to make a rational choice between various brands of a product. An advertisement by the Vanaspati Ghee Association, the Tea Board or the Coffee Board is of the informative type because it helps to push the sales of all firms in the group.

Competitive advertisement on the other hand, is meant to push the sales of the product of a particular firm as against other similar products.

Appealing posters, short films, commercial broadcasts showing a famous film star uttering words in praise of a particular product and requesting viewers and listeners to buy that product since it is the best, is persuasive or competitive advertisement. Its aim is to push the sales of one firm at the cost of others. We shall be concerned with this type of selling costs.

Of the two kinds of advertisements, the former is socially advantageous in all such cases where the products advertised are useful. Thus informative advertisement spreads knowledge. Competitive advertisement is, however, socially undesirable because it involves wastage of resources.

Production Costs vs. Selling Costs

Since each firm has to incur selling costs under monopolistic competition, its total costs include production costs and selling costs. Production costs include all expenses incurred in making a particular product, and transporting it to its destination for consumers.

They are the outlays on the services of all factor-services land, labour, capital and organisation—engaged in the manufacture of the product and also includes packaging, transport and service charges. Selling costs are costs incurred to change consumer preferences for a particular product. They are intended to raise the demand for one product rather than another at any given price.

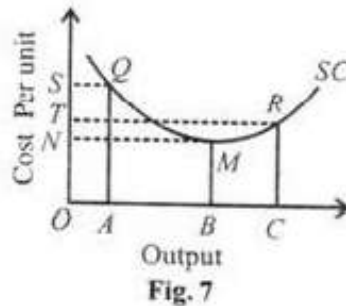
Prof. Chamberlin distinguishes between the two in these words: The former (production) costs create utilities in order that demands may be satisfied, the latter create and shift the demand curves themselves". To be precise, "those which alter the demand curve for a product are selling costs and those which do not, are production costs". In other words, "those made to adapt the product to the demand are production costs and those made to adapt the demand to the product are selling costs."

However, no clear-cut distinction can be made between production costs and selling costs. Is the cost of a cellophane wrapper, for instance, a production cost or a selling cost? In fact, "the two types of costs are interlaced throughout the price system, so that at no point such as at the competition or manufacture, can one be said to end and the other to begin". But Prof. Chamberlin himself finds a way out by confining selling costs to only advertising expenses.

The Curve of Selling Costs and Its Influence on Production Costs

The curve of selling costs is a tool of economic analysis forged by Prof. Chamberlin. It is a curve of average selling cost per unit of product. It is akin to the average cost curve and like the latter is U-shaped under the influence of the law of variable proportions.

The curve of selling costs first falls, reaches a minimum point and then starts rising.



SC is the curve of selling costs. AQ is the average cost of selling OA units of the product, the total cost of selling it being OAQS. At the minimum point M of the SC curve, the cost per unit of selling OB units is BM which is less than any point in the QM portion of the SC curve. After this point, the average selling cost of $\hat{N}$ units is RC, the total cost of selling this quantity of the product being $\hat{C}$ RT. In fact, per unit selling cost and total selling costs increase beyond the minimum point M.

According to Chamberlin, the shape of the curve and the exact point where it moves upward depends upon the nature of the product, its price, the nature of competing substitutes, the incomes of the buyers and their reluctance to change their tastes by the advertisement. There is, however, a limit to the rising portion of the selling cost curve. When the sales reach the saturation point, it ultimately becomes vertical.

In the beginning, the application of successive doses of selling costs will raise the total sales more than proportionately so that average selling costs fall. This is due to two factors. Firstly, consumers being attached to a particular brand of the product say, Brooke Bond Tea, are in the habit of buying it alone.

Advertisement in favour of another variety of the same product say Tata tea, is meant to break their habit and dissolve the attachment of consumers to Brooke Bond. One or two insertions a month in newspapers may make little impact on the consumers.

To convert the buyers to its brand, the manufacturer will have to incur larger selling outlays in the form of repetitive insertions of advertisement in newspapers, in commercial broadcasts, on the TV, in the form of samples, gifts or premium coupons.

Then only, sales will increase. Secondly, as larger outlays are incurred on sales promotion, internal economies of advertisement appear in the form of efficient salesmen, attractive advertisements and packing, etc. For instance, the larger the insertions and the size of advertisement, the lower the advertising rates per page. Thus, these two factors tend to lower average selling costs per unit of product up to a point.

Beyond this critical output M, average selling costs start rising again due to two forces. One, progressively increasing sales promotional expenses is to be incurred to induce regular consumers to continue to buy it. Efforts to induce old customers to buy the same product require larger promotional expenses so that they are not only dissuaded from buying some other brand but also persuaded to buy more of the same product.

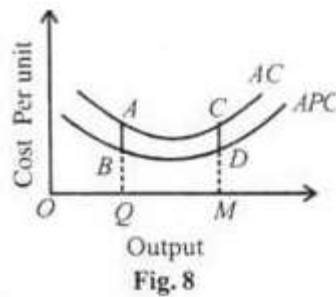
Two, larger selling outlays are required to attract new customers and those attached to other brands of the same product. They have to be convinced by repetitive advertisements in newspapers, through the radio, the cinema and the television, of the superiority of this particular brand. Naturally such efforts require increased selling expenses. As a consequence of these forces, average selling costs per unit of the product rise.

We may conclude that two sets of forces operate in response to selling outlays on a product which tend to bring increasing returns upto a point, and beyond that, diminishing returns.

Proportional Selling Costs

Average selling costs have the effect of raising the average total cost of production. If the average selling costs are proportional to the product sold, the curve of average total costs will lie at an equal distance above the average product costs curve. For instance, when with a Signal toothpaste, a packet of five Erasmic blades is given free the cost of five Erasmic blades incurred by the makers

of the toothpaste represents proportional selling costs. The product cost per unit of toothpaste and the selling cost per unit of a packet of blades added up from the total cost per unit of toothpaste-cum-five blades. The average production costs will rise by the cost of blades and will remain the same so long as the firm continues to sell in this proportion.



APC represents the average production costs curve and AC the average costs curve. The average costs are proportional throughout. They remain the same at all levels of output. At OQ output they are BA and even at OM level of output, they are the same as before $dc (=BA)$, so is the average total cost $mc (=QA)$. It should be noted that the corresponding mc curves of APC and ac will also move in the same proportion.

We have assumed here that the producer continues to incur proportional selling costs which are unrealistic. In fact, a producer will resort to proportional selling costs only for a short while till his old stock is exhausted and in the process he also attracts new customers and induces regular users to buy more of it.

Fixed Selling Costs

Selling costs are also of a fixed type, as in the case of screening a short-film in a cinema for a month or an insertion in the newspaper each Sunday. The average total cost per unit will at first be higher and as output increases it will fall and then, after a point, will start rising and the average total cost curve gradually becomes closer to the average production cost curve, as output increases further.

It implies that as sales increase, fixed selling costs are spread over a larger output and become less and less. Where APC

represents the average production cost curve and ac the average total cost curve inclusive of selling costs. The corresponding MC curves to APC and ac would be derived in the same manner and bear the same relation to their respective average cost curves. Added together they would give a combined mc curve.

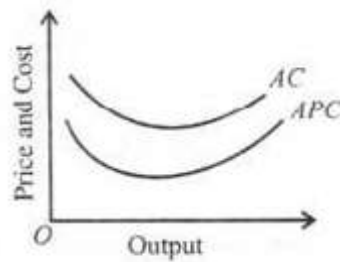


Fig. 9

Sometimes, certain tailoring and dry cleaning concerns provide free home-delivery service to their customers. But in such cases the effect of home-delivery service is difficult to calculate on average and marginal costs.

Influence of Selling Costs on the Demand Curve

The purpose of selling costs is to influence the demand curve for the product of a firm or group. A producer incurs selling costs in order to push up his sales. Therefore, all selling costs tend to shift an individual seller's demand curve to the right. The question of a demand curve shifting to the left is altogether ruled out in this analysis.

When the demand curve for the product of a firm shifts to the right, it is the result either of inducing the same customers to buy more of the same product or new customers buying this product attracted by the advertisement. The new demand curve may be more or less elastic throughout its length or parts of it may be more or less elastic than the old demand curve before incurring selling costs.

If the buyers are convinced of the superiority of this product in contrast with other similar products, the new demand curve will be less elastic in the upper segments than the old demand

curve. The firm will be losing few customers as a result of rise in the price of its product. If, on the other hand, the old and the new customers are attracted more to the product, but, at the same time, are not prepared to pay a very high price, the new demand curve will be highly elastic in its lower portion.

Besides, the buying habits of the old and new customers also influence the shape of the demand curve. If they are influenced more by price changes rather than the product variation, the new demand curve will be highly elastic. On the contrary, if they are not influenced much by price changes, the new demand curve will be less elastic than the old curve.

In the analysis that follows straight line demand curves are taken for the sake of simplicity. When the new demand curve is drawn parallel to the old, the elasticity of demand on the higher curve is lower at each price. It means that the consumers are convinced of the superiority of this product and are willing to pay a higher price.

Price-Output Determination under Selling Costs

Under monopolistic competition, an individual firm has a variety of choices to sell a larger output. It may do so by lowering the price of the product; it may improve its quality; it may indulge in greater sales promotion efforts; or it may resort to all the three methods simultaneously or combine the one with the other. We shall, however, be concerned with only selling costs.

But even this problem is complicated because to depict each level of output at each price and the possible AR, MR, MC and AC curves on a two-dimensional figure becomes complex. So for the sake of simplicity, the demand curve and the average total cost curve are taken along with the average production cost curve. The following analysis discusses the influence of the selling costs on the price-output policy of the firm.

COMPARISON WITH PERFECT COMPETITION AND MONOPOLY

Perfect competition is often distinguished from pure competition, but they differ only in degree. The first five conditions

relate to pure competition while the remaining four conditions are also required for the existence of perfect competition. According to Chamberlin, pure competition means, competition unalloyed with monopoly elements," whereas perfect competition involves perfection in many other respects than in the absence of monopoly."

The practical importance of perfect competition is not much in the present times for few markets are perfectly competitive except those for staple food products and raw materials. That is why, Chamberlin says that perfect competition is a rare phenomenon."

Though the real world does not fulfil the conditions of perfect competition, yet perfect competition is studied for the simple reason that it helps us in understanding the working of an economy, where competitive behaviour leads to the best allocation of resources and the most efficient organisation of production. A hypothetical model of a perfectly competitive industry provides the basis for appraising the actual working of economic institutions and organisations in any economy.

Competition and Monopoly

Types of monopoly-price: receipts vs. profits. These petty devices develop, in the case of larger markets and of many important articles of sale, into the systematic practice of manipulating prices artificially.

The explanation of the motives and of the limits of monopolistic price-fixing would best be reserved in large part until a later stage of our study, where it can be considered in connection with enterprise. It is in the sale of the products of a business that the most important problems of monopoly are found. There the monopolist is seeking the highest net gain over a considerable period in the sale of a continuous output of goods.

The cost per unit is the minimum seller's valuation and the monopoly-price sought is that which in the long run yields the largest gain (the product of units of sales times margin of gain per unit). Let us here consider merely the case where the monopoly (seller or group of sellers) is seeking the maximum total price (not

net gain) for a stock of goods which have no minimum seller's valuation.

Such is the classic example of monopoly in colonial trade related by Adam Smith: "In the spice islands the Dutch are said to burn all the spiceries which a fertile season produces beyond what they expect to dispose of in Europe with such a profit as they think sufficient." "We may call this price which concerns the gross receipts from sales, crude monopoly-price. It is that which yields the monopolist (with complete control of supply) the maximum gross receipts.

This type of cases is of not infrequent occurrence. Such a case is presented whenever the unsold portion of a supply would go to waste, such as perishable goods after they have come to market (fruits, vegetables, etc.), such as vacant seats in an opera house, at athletic games, etc., where the expense of the whole performance has been incurred and will not be increased by more spectators. This control of all the seats at a single entertainment is a very restricted kind of monopoly, and does not present a social problem. There is still intense competition among artists of all kinds to provide entertainments having the merits to attract spectators.

Uniform Monopoly-price

In all such cases the competitive price would be fixed solely by the buyers' scale of valuation, as in an auction without reserve. If the supply of goods be large, approaching the saturation point of desires, whether there be one seller (without reserve valuation) or competing sellers, the price will tend toward the valuation of the marginal buyer, and in the extreme case may sink to zero. The only way sellers can prevent this is to reserve a part of the supply, even if it has to be burned up or thrown away (fish, fruit, etc.), or remains unused (as the empty seats in a theater). In the case shown if there were 7 units for sale, the unit price would be 1, the total price 7, and each of the 7 sellers would get 1.

But if the owners of these 7 units unite and withhold 3 units, the total receipts are 16, which divided equally, gives $22/7$ units of price to each seller. It is a general truth, that monopoly power can

be made effective to raise a uniform market-price above what it would be if the monopolistic competed, only by artificially increasing scarcity, by limiting supply. Shown graphically, the maximum crude monopoly-price obtainable is always the largest rectangle that can be inscribed within the coordinate axes and the hypothetical demand-curve.

Uniform Monopoly-price, Inelastic Demand

With a more inelastic demand, where buyers' demand increases very little with a rapid fall in price, the monopolist must restrict his offers more narrowly to attain a total price above the competitive. The offer of 3 units would at the price 6 yield the maximum proceeds (18), and any supply below that would be tapping only the lower levels of valuation. If a few valuations are high, and the others fall very rapidly, the price can be raised very much more, if the demand curve were AEFG the monopoly-price would be 9. This is the type of demand for articles of great luxury, limited to the very rich. Note that as demand means number of units demanded, at a price, an elastic demand means a large change of demand with a small change in price. With given scales of price and of quantities of goods, the more elastic the demand, the flatter the demand-curve.

Uniform monopoly-price: elastic demand. The more elastic the demand the more nearly a monopolistic price approaches a competitive price with a given number of units of supply. It appears that with any number of units up to 4, the monopolistic and the competitive market-prices would be the same, and any restriction would involve a loss to the monopolist.

The motive for monopoly lies in the range of supply of 5 units and beyond. With a more elastic type of demand as in scale A-B where there is less difference in the valuations of the most urgent (or capable) and of the less urgent buyer, competitive and monopolistic market-prices are the same up to 7 units ($7 \times 41/3 = 301/3$ total). With a still more elastic demand represented by a more flattened curve, as in C-D, the competitive and monopolistic price are the same up to 11 units ($11 \times 31/3 = 36\%$ total) and either 10 or 11 units will yield the same total. Beyond that is the region of

possible monopolistic price. Compound types of demand scales, made up of different levels of demand, would further strengthen or weaken the motive to limit supply. If the demand curve, after rapidly falling, flattens to a new broad field of demand, a lower price will yield a larger total than the previous monopoly-price. This is the type of non-essential goods which remain luxuries when price is high, but rapidly become looked upon as comforts and necessities when price falls.

Discriminatory monopolistic price. It appears from the foregoing that while it is possible for sellers to gain by the fixing of a uniform monopoly-price under some conditions, under other cases it is not. The range of this possibility is, indeed, much narrower than would be anticipated before a study of the problem. But where a monopoly exists, why should it confine itself to a uniform price to all buyers? The very scrutiny of the differences in buyers' valuations needed to fix a monopoly-price, suggests making differences in prices. This fact of practical experience presents the problem of discriminatory monopoly-price.

It may often happen that the whole group of would-be buyers may be divided into subgroups, and a different price made for each. This division may correspond with differences in locality (geographical), as in railroad rates to different places, different prices of petroleum to different cities or states, or different rates to domestic and to foreign shippers on a railroad, etc. Or it may correspond with social ranks, as can be done by making slight differences in quality, the best quality at a very high price for the rich, and the common grades at low prices to the masses.

Or it may correspond with the power of different buyers to substitute other goods, or to resort to a different source of supply, the poor in such cases being made to pay more than the rich. Or the distinction may be made with reference to the individual differences in maximum valuations. It must not be forgotten that our study thus far is limited to crude monopoly-price. The problem is different when it is one of profits resulting from the excess of price over cost of production only to be known by intimate personal knowledge or by an elaborate system of espionage. This is the extremest possible discrimination.

OLIGOPOLY: CHARACTERISTICS, PRICE LEADERSHIP, KINKED DEMAND CURVE

Oligopoly is a market situation in which there are a few firms selling homogeneous or differentiated products. It is difficult to pinpoint the number of firms in 'competition among the few.' With only a few firms in the market, the action of one firm is likely to affect the others. An oligopoly industry produces either a homogeneous product or heterogeneous products.

The former is called pure or perfect oligopoly and the latter is called imperfect or differentiated oligopoly. Pure oligopoly is found primarily among producers of such industrial products as aluminium, cement, copper, steel, zinc, etc. Imperfect oligopoly is found among producers of such consumer goods as automobiles, cigarettes, soaps and detergents, TVs, rubber tyres, refrigerators, typewriters, etc.

Characteristics of Oligopoly

In addition to fewness of sellers, most oligopolistic industries have several common characteristics which are explained below:

Interdependence

There is recognised interdependence among the sellers in the oligopolistic market. Each oligopolist firm knows that changes in its price, advertising, product characteristics, etc. may lead to counter-moves by rivals. When the sellers are a few, each produces a considerable fraction of the total output of the industry and can have a noticeable effect on market conditions.

He can reduce or increase the price for the whole oligopolist market by selling more quantity or less and affect the profits of the other sellers. It implies that each seller is aware of the price-moves of the other sellers and their impact on his profit and of the influence of his price-move on the actions of rivals. Thus there is complete interdependence among the sellers with regard to their price-output policies. Each seller has direct and ascertainable influences upon every other seller in the industry. Thus, every move by one seller leads to counter-moves by the others.

Advertisement

The main reason for this mutual interdependence in decision making is that one producer's fortunes are dependent on the policies and fortunes of the other producers in the industry. It is for this reason that oligopolist firms spend much on advertisement and customer services.

As pointed out by Prof. Baumol, "Under oligopoly advertising can become a life-and-death matter." For example, if all oligopolists continue to spend a lot on advertising their products and one seller does not match up with them he will find his customers gradually going in for his rival's product. If, on the other hand, one oligopolist advertises his product, others have to follow him to keep up their sales.

Competition

This leads to another feature of the oligopolistic market, the presence of competition. Since under oligopoly, there are a few sellers, a move by one seller immediately affects the rivals. So each seller is always on the alert and keeps a close watch over the moves of its rivals in order to have a counter-move. This is true competition.

Barriers to Entry of Firms

As there is keen competition in an oligopolistic industry, there are no barriers to entry into or exit from it. However, in the long run, there are some types of barriers to entry which tend to restrain new firms from entering the industry.

They may be: (a) Economies of scale enjoyed by a few large firms; (b) control over essential and specialised inputs; (c) high capital requirements due to plant costs, advertising costs, etc. (d) exclusive patents and licenses; and (e) the existence of unused capacity which makes the industry unattractive. When entry is restricted or blocked by such natural and artificial barriers, the oligopolistic industry can earn long-run super normal profits.

Lack of Uniformity

Another feature of oligopoly market is the lack of uniformity

in the size of firms. Firms differ considerably in size. Some may be small, others very large. Such a situation is asymmetrical. This is very common in the American economy. A symmetrical situation with firms of a uniform size is rare.

Demand Curve

It is not easy to trace the demand curve for the product of an oligopolist. Since under oligopoly the exact behaviour pattern of a producer cannot be ascertained with certainty, his demand curve cannot be drawn accurately, and with definiteness. How does an individual seller's demand curve look like in oligopoly is most uncertain because a seller's price or output moves lead to unpredictable reactions on price-output policies of his rivals, which may have further repercussions on his price and output.

The chain of action reaction as a result of an initial change in price or output, is all a guess-work. Thus a complex system of crossed conjectures emerges as a result of the interdependence among the rival oligopolists which is the main cause of the indeterminateness of the demand curve.

If the oligopolist seller does not have a definite demand curve for his product, then how does he affect his sales. Presumably, his sales depend upon his current price and those of his rivals. However, a number of conjectural demand curves can be imagined.

For example, in differentiated oligopoly where each seller fixes a separate price for his product, a reduction in price by one seller may lead to an equivalent, more, less or no price reduction by rival sellers. In each case, a demand curve can be drawn by the seller within the range of competitive and monopoly demand curves.

Leaving aside retaliatory price movements, the individual seller's demand curve under oligopoly for both price cuts and increases is neither more elastic than under perfect or monopolistic competition nor less elastic than under monopoly. It may still be indefinite and indeterminate.

KD_1 is the elastic demand curve and MD is the less elastic demand curve. The oligopolies' demand curve is the dotted kinked

KPD. The reason is quite simple. If a seller reduces the price of his product, his rivals also lower the prices of their products so that he is not able to increase his sales.

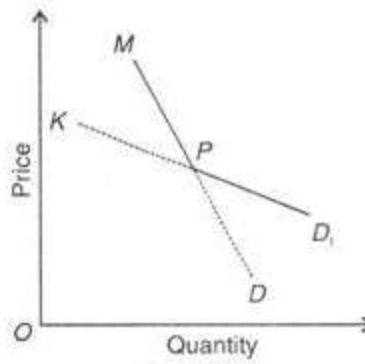


Fig. 1.

So the demand curve for the individual seller's product will be less elastic just below the present price P (where KD_1 and MD curves are shown to intersect). On the other hand, when he raises the price of his product, the other sellers will not follow him in order to earn larger profits at the old price. So this individual seller will experience a sharp fall in the demand for his product.

Thus his demand curve above the price P in the segment KP will be highly elastic. Thus the imagined demand curve of an oligopolist has a corner or kink at the current price P . Such a demand curve is much more elastic for price increases than for price decreases.

No Unique Pattern of Pricing Behaviour

The rivalry arising from interdependence among the oligopolists leads to two conflicting motives. Each wants to remain independent and to get the maximum possible profit. Towards this end, they act and react on the price-output movements of one another in a continuous element of uncertainty.

On the other hand, again motivated by profit maximisation each seller wishes to cooperate with his rivals to reduce or eliminate the element of uncertainty. All rivals enter into a tacit or formal

agreement with regard to price-output changes. It leads to a sort of monopoly within oligopoly.

They may even recognise one seller as a leader at whose initiative all the other sellers raise or lower the price. In this case, the individual seller's demand curve is a part of the industry demand curve, having the elasticity of the latter. Given these conflicting attitudes, it is not possible to predict any unique pattern of pricing behaviour in oligopoly markets.

Demand Curve Shifts

When consumers increase the quantity demanded *at a given price*, it is referred to as an *increase in demand*. Increased demand can be represented on the graph as the curve being shifted to the right. At each price point, a greater quantity is demanded, as from the initial curve D1 to the new curve D2.

In the diagram, this raises the equilibrium price from P1 to the higher P2. This raises the equilibrium quantity from Q1 to the higher Q2. A movement along the curve is described as a "change in the quantity demanded" to distinguish it from a "change in demand," that is, a shift of the curve.

In the example above, there has been an *increase* in demand which has caused an increase in (equilibrium) quantity. The increase in demand could also come from changing tastes and fashions, incomes, price changes in complementary and substitute goods, market expectations, and number of buyers. This would cause the entire demand curve to shift changing the equilibrium price and quantity.

If the *demand decreases*, then the opposite happens: a shift of the curve to the left. If the demand starts at D2, and *decreases* to D1, the price will decrease, and the quantity will decrease. This is an effect of demand changing. The quantity supplied at each price is the same as before the demand shift (at both Q1 and Q2). The equilibrium quantity, price and demand are different. At each point, a greater amount is demanded (when there is a shift from D1 to D2). The demand curve "shifts" because a non-price determinant of demand has changed. Graphically the shift is due

to a change in the x-intercept. A shift in the demand curve due to a change in a non-price determinant of demand will result in the market's being in a non-equilibrium state. If the demand curve shifts out the result will be a shortage — at the new market price quantity demanded will exceed quantity supplied.

If the demand curve shifts in, there will be a surplus — at the new market price quantity supplied will exceed quantity demanded. The process by which a new equilibrium is established is not the province of comparative statics — the answers to issues concerning when, whether and how a new equilibrium will be established are issues that are addressed by stochastic models — economic dynamics.

Factor Pricing

Price is a matter of vital importance to both the seller and the buyer in the market place. In money economy, without prices there cannot be marketing. Price denotes the value of a product or service expressed in money. Only when a buyer and a seller agree on price, we can have exchange of goods and services leading to transfer of ownership.

In a competitive market economy, price is determined by free play of demand and supply. The price will move forward or backward with changing supply and demand conditions. The going market price acts as basis for fixing the sale price. Rarely an individual seller can dishonour the current market price. In a free market economy, we have freedom of contract, freedom of enterprise, free competition and right to private property.

Price regulates business profits, allocates the economic resources for optimum production and distribution. Thus, price is the prime regulator of production, distribution and consumption of goods. Economics revolves around pricing of resources. Price influences consumer purchase decisions. It reflects purchasing power of currency. It can determine the general living standards. In essence, by and large, every facet of our economic life is directly or indirectly governed by pricing. This is literally true in our money and credit economy. Pricing decisions interconnect marketing actions with the financial objectives of the enterprise. Among the most important marketing variables influenced by pricing decisions are: 1. sales volume, 2. profit margins, 3. rate of return on investment, 4. trade margins, 5. advertising and sales

promotion, 6. product image, 7. new product development. Therefore, pricing decisions play a very important role in the design of the marketing mix. Pricing strategy determines the firm's position in the market vis-avis its rivals. Marketing effectiveness of pricing policy and strategy should not suffer merely on account of cost and financial criteria. Price is a powerful marketing instrument. As a marketing weapon, pricing is the *big-gun*. However, it must be used with great caution. It is a dangerous and explosive marketing force. It may doom a good product to failure. Low pricing strategies are irreversible decisions. They must be used correctly from the outset. Every marketing plan involves a pricing decision. Therefore, all marketing planners must make accurate and planned pricing decisions.

NEED OF SEPARATE THEORY OF FACTOR PRICING

The Significance of the Price Factor

The selling price plays a unique role in business because the price level:

1. controls the sales volume and the firm's market share, 2. determines the total sales revenue (sales revenue=sales volume X unit price), 3. regulates the rate of return on investment (ROI) and through ROI price influences sales profitability, 4. creates an impact on unit cost in mass production. Low price increases total production and sales turnover, and ultimately mass production (through economies of scale) leads to the lower unit cost of production. Low price induces also efficiency in production and marketing. Henry Ford stated: "Our policy is to reduce the price, extend operations and improve the product.

Comments:

1. All other elements (except price) in the marketing mix are called non-price factors. They influence price and are also influenced by price. All elements are interdependent interacting factors.
2. We have two relationships: (a) Cost/Price relationship, and (b) Price/Revenue relationship.

3. Price and other marketing mix variables are complementary factors. They may be partial substitutes for each other.
4. Together all elements in the marketing mix collaborate to accomplish a common objective, *viz.*, to produce sales and sales revenues.

Typical Pricing Objectives

A variety of objectives may guide pricing decision:

1. **Growth in Sales:** A low price can achieve the objective of increase in sales volume. A low price is not always necessary. A right price can stimulate the desired sales increase. In practice, price and non-price objectives are coordinated to produce the desired increase in sales. Competitive price, if used wisely, can secure faster increase in sales than any other marketing weapon.
2. **Market Share:** Price is typically one of those factors that carries the heaviest responsibility for improving or maintaining market share — a sensitive indicator of customer and trade acceptance.
3. **Predetermined Profit Level:** Return on Investment, say 20 to 25 per cent is a common decision in marketing. Pricing for profit is the most logical of all pricing objectives.
4. **Meet or Follow Competition:** Many firms desire the stabilisation of price levels and operating margins as more important than the maintenance of a certain level of short-run profits. The price leader maintains stable prices in the industry. Follow the leader.

Control Cash-flow: A principal pricing objective is to return cash as much as possible (the funds invested) within a given period. Investment in research and development, market development, promotion, etc., should pay back within a specified period. Capital expenditure on any project must be recovered within 5 to 10 years. Pay-back or cash-flow objectives fits in easily with other corporate objectives.

Note: 1. While determining objectives of a pricing policy, marketers must take into account reactions of a number parties

such as customers, competition, resellers or dealers, government, public opinion, and so on. 2. The objectives may not be mutually exclusive. Marketers have to resolve their conflicts. For instance, there may be a conflict between sales maximisation objective and a return on investment or profit objective. However, it should be noted that maximum market penetration in the short run (in the early phase of the product life cycle) is the key to maximum ROI in the long run.

Market Price

The market price is the price determined by the free play of demand and supply. The market price of a product affects the price paid to the factors of production — rent for land, wages for labour, interest for capital and profit for enterprise. In this way, price becomes a prime or basic regulator of the entire economic system because it influences the allocation (distribution) of these resources (factors of production). For example, when the price of a commodity has a rising tendency, we shall have higher wages attracting more labour, higher interest attracting more capital, and so on, in the industry in which prices are rising. Conversely, under falling prices, low wages, low rent, low interest, and low profits will reduce the availability of labour, land, capital and risk-takers in a free market economy. Prices direct and control production and consumption.

Since market price is determined in an impersonal way through the general relations of demand and supply, the individual seller has no control over the market price and the actual market price at any given time may be above or below the costs of individual sellers. Market price is indicated by published prices, market reports, etc. A seller will have to change his output to adjust with the current market price in order to secure maximum gains or minimise his losses. He can also minimise operating costs.

Price as a Measure of Value

Economic theory of price has a few simple assumptions regarding products and buyer behaviour. Buyer's tastes and preferences are considered as given (constant). Buyer is considered essentially a rational human being. The marketing concepts like

brand image, brand loyalty and benefit segmentation 'emotional motivation' are outside the scope of price theory. Hence, in practice, the classical price theory, saying price determines value of the product, is not true.

Marketers have recognised the importance of perception, learning, and attitudes creating psychological reactions to price, at least in consumer goods. The social and psychological factors must be recognised in the evaluation of pricing strategies. The social and psychological influences are responsible to support the consumer's inclination to use price as an indicator of quality for certain products *e.g.*, cosmetics, jewellery, and clothing. Such products have concealed values and benefits which the consumer cannot evaluate rationally or on objective basis.

Consumer does not have physical cues or guides suggesting product quality in many cases and social psychological dimensions may dominate in the consumer behaviour. Under such situations, price is the most handy (but rough) available indicator of product quality and value for many customers.

Buyers believe in the implicit subjective process *viz.*, "You get what you pay for", "If it costs more, it must be better". Marketers are bound to exploit buyer's emotions, preferences and habits. Price-quality relationship applies to products whose quality is difficult to judge and whose brands vary widely in quality, *e.g.*, cosmetics jewellery, clothing, wine, floor wax, etc.

Charm pricing is another psychological dimension of pricing. Accepted pricing conventions have a charm for the consumer, *e.g.*, price like Rs. 99. The quotation of Rs. 19.90 sounds better value than Rs. 20. Price lining is another psychological dimension of pricing accounting for a common marketing practice. For example, a reasonable price range for a new Television set is between Rs. 8,000/- and Rs. 14,000/- for most people. Only handful of buyers would seriously consider purchasing TV set costing Rs. 20,000/- or more, and new TV set costing less than Rs. 4,000/- would generate doubts and suspicion.

The consumers answer the question (Is it worth it?) in terms of the familiar equation:

Satisfaction = Benefit — Cost

The price is the cost part of the equation. It indicates sacrifice of purchasing power.

The reasons for the inability of price to determine the perceived value of the product are:

There are considerable differences in the market information available to consumers. 2. We have significant differences in the bargaining power of consumers. 3. In large parts of the retail market we have non-price competition replacing price competition. The purpose of non-price competition is also to make sales or demand curve less sensitive to price and the price of an article might be raised without adverse effect on sale (demand has become less price elastic due to promotion). The higher price compensates for promotion costs incurred in stimulating demand.

MARGINAL PRODUCTIVITY THEORY OF FACTOR PRICING**Marginal Productivity Theory**

This theory was propounded by Phillips Henry Wick-steed (England) and John Bates Clark of U.S.A. According to this theory, wages is determined based on the production contributed by the last worker, i.e. marginal worker. His/her production is called 'marginal production'.

The Bargaining Theory of Wages

John Davidson was the propounder of this theory. According to this theory, the fixation of wages depends on the bargaining power of workers/trade unions and of employers. If workers are stronger in bargaining process, then wages tends to be high. In case, employer plays a stronger role, then wages tends to be low.

Behavioural Theories of Wages

Based on research studies and action programmes conducted, some behavioural scientists have also developed theories of wages. Their theories are based on elements like employee's acceptance to a wage level, the prevalent internal wage structure, employee's consideration on money or wages and salaries as motivators.

CONCEPTS: RENT, WAGE, INTEREST AND PROFIT

Law of rent

The law of rent was formulated by David Ricardo around 1809, and presented in its most developed form in his magnum opus, *On the Principles of Political Economy and Taxation*. This is the origin of the term Ricardian rent. Ricardo's formulation of the law was the first clear exposition of the source and magnitude of rent, and is among the most important and firmly established principles of economics.

Description

The law of rent states that the rent of a land site is equal to the economic advantage obtained by using the site in its most productive use, relative to the advantage obtained by using marginal (i.e., the best rent-free) land for the same purpose, given the same inputs of labor and capital.

Ricardo formulated this law based on the principles put forth by Adam Smith in *Wealth of Nations*.

"The rent of land, therefore, considered as the price paid for the use of the land, is naturally a monopoly price. It is not at all proportioned to what the landlord may have laid out upon the improvement of the land, or to what he can afford to take; but to what the farmer can afford to give." — Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Book I, Chapter XI "Of the Rent of Land"

Ricardian rent should not be confused with contract rent, which is the "actual payments tenants make for use of the properties of others." (Barlow 1986). Rather, the law of rent refers to the economic return that land should accrue for its use in production.

Being a political economist, Ricardo was not simply referring to land in terms of soil. He was primarily interested in the economic rent and locational value associated with private appropriation of any natural factor of production. The law of rent applies equally well to urban land and rural land, as it is a fundamental principle of economics.

Ricardo noticed that the bargaining power of laborers can never dip below the produce obtainable on the best available rent-free land, because whenever rent leaves them with less than they could get on that free land, they can simply move to the new location. The produce obtainable on the best available rent-free land is known as the margin of production. Since landlords have a monopoly over a given location, the only limiting factor for rent is the margin of production. Thus, rent is a differential between the productive capacity of the land and the margin of production.

Note that Ricardo's original formulation assumes that the best quality land would be the first to be used in production, and that goods are sold in a competitive, single price market.

Ricardian Theory of Rent

The classical theory of rent is associated with the name of David Ricardo. He begins with a group of new settlers in a new country. Let us suppose ourselves to be the settlers in a hitherto unknown island which we shall call Jawahar Island after our late beloved leader.

As we study the natural resources of Jawahar Island, we find the land to be of four grades. For convenience, we call them A, B, C and D in the order of their fertility. We shall settle down in Tarapur in 'A' part of the island.

This is the most fertile land and gives us the largest produce per acre. Enough land is available of this quality to satisfy all our needs at the moment. Therefore, it is a free good and will not command any price, i.e., rent. But as time passes, the mouths to be fed increase in number. This may be due to more immigrants, who have heard of our good luck, or due to an increase in population.

Rent in Extensive Cultivation

A time comes when all land of the best quality has been taken up. But some demand still remains unsatisfied. We have then to resort to 'B' quality land.

It is inferior to 'A' and yields only 30 quintals of wheat per plot as compared with 35 quintals of 'A' with the same expenditure

of labour and capital. Naturally plots in 'A' now acquire a greater value as compared with 'B'. A tenant will be prepared to pay up to 5 quintals of wheat in order to get a plot in the 'A' zone, or take 'B' quality land free of charge.

This difference, paid to the owner (if the cultivator is a tenant) or kept to himself (if he is the owner), is economic rent. In the first case (i.e., when the cultivator is a tenant) it is contractual rent; and in the latter (i.e., when the cultivator is the owner) it is known as implicit rent. 'B' plots do not pay any rent. To go a step further, we see that after all land of 'B' quality has also been taken up, we begin cultivating 'C' plots. Now even 'B' quality land comes to have differential surplus over 'C'. Rent of 'A' increases still further.

When the demand increases still more, we are pushed to the use of the worst land, which is of 'D' quality yielding 25 quintals per plot. 'D' quality land is now no-rent land or marginal land while 'A', 'B', 'C' all earn rent. This growing demand shows itself in rising prices. They raise high enough to cover the expenses of cultivation on the lowest grade land, i.e., 'D' quality.

Let us suppose that one unit of productive effort is equal to Rs. 3,500. When only A' quality land, where a plot produces 35 quintals is under the plough, the price of wheat will be Rs. 100 per quintal. When owing to increased demand, the price of wheat rises to Rs. 110 then and only then will 'B' quality land be cultivated which produces 30 quintals of wheat. And when that happens 'A' land will have a surplus of 5 quintals $\times$ Rs. 110 = Rs. 550 per plot. This becomes rent.

The difference, in other words, between the return from a plot of land above the margin and the marginal plot (i.e., the one just paying its way) is called rent or economic rent.

Rent in Intensive Cultivation

The settlers in Jawahar Island realize that there is another way too of increasing the produce. Why not apply more labour and capital to superior lands, and resort to intensive cultivation? This is done but it is seen that the law of diminishing returns sets in. Now consider that A, B, C and D are the different doses of labour

and capital (instead of different grades of land) applied to the same grade of land. The first dose yields 35 quintals.

The second unit of labour and capital used on 'A' plot will almost definitely give us less than the first. We suppose it gives us only 30 quintals.

So we have the choice of either taking new plots in 'B' land, or cultivating 'A' lands more intensively. If we adopt the latter course, the first unit of labour and capital will be yielding a surplus over the second unit—which unit produces just enough to cover the expenses. This surplus, again, is rent. As more and more units of labour and capital are applied, the return per unit will go on falling.

Rent Due to Differential Advantages

With the passage of time, however, a new factor emerges. A locality in the A' zone—marked Tarapur—develops into a market and Azadnagar in 'B' into a railway junction, and produce has to be sent to those two flourishing localities for their final disposal. Now the plots situated in the neighbourhood of Tarapur and Azadnagar come to have an advantage. They have either no transport charges or much smaller charges than in the case of lands in 'C' and 'D' areas.

Transport charges are a part of the cost of production, because production is complete only when the commodity reaches the hands of consumers. The better-situated plots, which have to bear less transport charges, will enjoy a surplus over the distant ones. This surplus will be another cause of rent. Hence, economic rent is a surplus which arises on account of natural differential advantages, whether of fertility or of situation, possessed by the land in question over the marginal land.

No-rent or Marginal Land

The cases described above show that rent is earned due to a certain piece being better suited for cultivation or being better situated in regard to markets. But better than what? Of course better than some other plot of land. This 'some other' plot is marginal land which just covers its expenses and no more. This

land is called 'no-rent land'. All rents are measured from it upwards. It is quite possible that we may not be able to spot the 'no-rent land' because:

- (a) It may be paying scarcity rent, or
- (b) The owner might have invested some capital in it and the interest thereon might be mistaken for rent, or
- (c) The no-rent land may be in some other country or
- (d) The no-rent tracts may form part of a rent-paying area and be concealed in it.

Scarcity Rent

In our new home-country, Jawahar Island, we at last come to a situation when all the lands have been brought under the plough, and are being cultivated intensively too. But the price rises still further under the pressure of demand. Population has been increasing fast. Our country has become old and no more land is available as we are an island country. Prices of agricultural produce go up and, therefore, incomes from land go up.

Hence, all land (including the no-rent 'D' quality land) begins to get surplus above expenses. This surplus above costs in the 'D' quality land, our previous no-rent land, is scarcity rent. Superior lands will be paying this surplus over and above differential gain.

Conclusion: Summing up, we can say that, according to the Ricardian theory, rent is a differential surplus and arises from the fact that land possesses certain peculiarities as a factor of production. It is limited in area and its fertility varies. Besides, its situation is fixed.

Thus rent results because:

- (a) Fertility is more or less fixed by nature;
- (b) The total stock of land is fixed and cannot be increased.

On this basis, Ricardo defines rent as "that portion of the produce of the earth which is paid to the landlord for the original and indestructible powers of the soil." According to him fertility, situation and limited total stock – these qualities, which are original as well as permanent, give, rise to rent.

Wage theory

The subsistence theory of wages, advanced by David Ricardo and other classical economists, was based on the population theory of Thomas Malthus. It held that the market price of labour would always tend toward the minimum required for subsistence. If the supply of labour increased, wages would fall, eventually causing a decrease in the labour supply. If the wage rose above the subsistence level, population would increase until the larger labour force would again force wages down.

The wage-fund theory held that wages depended on the relative amounts of capital available for the payment of workers and the size of the labour force. Wages increase only with an increase in capital or a decrease in the number of workers. Although the size of the wage fund could change over time, at any given moment it was fixed. Thus, legislation to raise wages would be unsuccessful, since there was only a fixed fund to draw on.

Karl Marx, an advocate of the labour theory of value, believed that wages were held at the subsistence level by the existence of a large number of unemployed.

The residual-claimant theory of wages, originated by the American economist Francis A. Walker, held that wages were the remainder of total industrial revenue after rent, interest, and profit (which were independently determined) were deducted.

In the bargaining theory of wages, there is no single economic principle or force governing wages. Instead, wages and other working conditions are determined by workers, employers, and unions, who determine these conditions by negotiation.

The marginal productivity theory of wages, formulated in the late 19th century, holds that employers will hire workers of a particular type until the addition to total output made by the last, or marginal, worker to be hired equals the cost of hiring one more worker. The wage rate will equal the value of the marginal product of the last-hired worker.

Supporters of this theory maintain that the test of an economic theory should be its predictive power. They hold that the marginal-productivity theory is a guide to long-run trends in wage

determination and applies more generally than the bargaining theory of wages.

Marginal Productivity Theory of Wages: Statement and Limitations of the Theory

Statement of the Theory

The marginal productivity theory states that, under conditions of perfect competition, every worker of same skill and efficiency in a given category will receive a wage equal to the value of the marginal product of that type of labour.

The marginal product of labour in any industry is the amount by which the output would be increased if one more man was employed while the quantities of other factors of production employed in the industry remained constant.

In short, it is the output of a single worker unaccompanied by any change in other factors of production.

The value of the marginal product of labour is the price at which the marginal product can be sold in the market. Under conditions of perfect competition, an employer will go on employing more and more workers until the value of the product of the last man he employs is equal to the marginal or additional cost of employing the last man.

Further, the condition of perfect competition implies that the marginal cost of labour is always equal to the wage rate, irrespective of the number of men the employer may engage. Every industry being ultimately subject to law of diminishing returns, this marginal product must start declining sooner or later. Wages remaining the same, the employer stops employing more workers at that point where the value of the product of a worker is equal to the wage rate.

So far we have assumed that the quantities of other factors remain constant while that of labour alone increases. This, however, is not realistic, because quantities of other factors too can be increased, though this may not be true in the short run.

To allow for this fact, the economists make use of the term “marginal net product of labour” instead of “marginal product of

labour". The value of marginal net product of labour may be defined as being the value of the amount by which output would be increased by employing one more man with the appropriate addition of other factors of production, less the addition to the cost of the other factors caused by increasing the quantities of other factors.

The theory may thus finally be re-stated as follows:

Under conditions of perfect competition in the labour market and in the market for the products of the industry, and irrespective of the number employed, every worker will receive a wage equal to the value of marginal net product of his labour.

Limitations of the Theory

We have already studied in detail the various limitations and criticisms of the Marginal Productivity Theory as a general principle of distribution. With reference to its application to wages, we may repeat that the theory is true only under certain assumptions such as perfect competition, perfect mobility of labour from employment to employment, homogeneous character of all labour, constant rates of interest and rent and given prices of the product. It is a static theory.

The actual world is dynamic. All the factors assumed to be constant are in fact constantly changing, competition is never perfect; mobility of labour is restricted for various reasons; all labour is not of the same grade, remuneration to other factors of production does not remain constant; and the prices of the products of labour vary. All these changes modify the theory when applied to actual conditions. The theory, however, as an assertion of a tendency is true and is valuable in understanding the basic forces that determine wage rates.

In the actual world, owing to the absence of the above assumptions, there is no single rate of wages that may be applicable to all labour of a particular type. Wages differ from place to place, from person to person and from employment to employment.

The following limitations or points of criticism of the marginal productivity theory may now be noted:

Firstly, this theory has little applicability to reality: The labour is not perfectly mobile. Workers of the same skill and efficiency may not receive the same wages at two different places.

Secondly: Though the condition of a large number of independent sellers is fulfilled for a few industries of all countries and for most industries of some countries, the employers usually combine to the disadvantage of the workers. It is a case of monopsony, i.e., one buyer (i.e., the employer) and many sellers (i.e., workers). The employers succeed in pulling down the wages below the value of the marginal net product of labour.

If employees are also collectively organized, the wage rates may or may not be equal to the values of marginal net product of labour in the occupations or industries concerned. The wages are determined by the relative bargaining strength of the two parties, but will not for a long time exceed the value of the marginal net product of labour.

Thirdly: The market for goods is in general characterised by imperfect competition. This also upsets the theory.

Fourthly: The productivity of workers is also dependent on factors such as the quality of capital and efficiency of management. These factors are beyond the control of workers.

Fifthly: Productivity is also dependent on wages. Low productivity may be the cause of low wages, which may tell on the efficiency of the worker, lower his standard of living, and ultimately check the supply of labour. The theory takes the supply of labour for granted.

In short, the marginal productivity theory ignores the effect of wage changes on the supply of labour, bargaining strength and monopoly conditions, etc.

The Classical Theory of Interest

The classical theory of interest also known as the demand and supply theory was propounded by the economists like Marshall and Fisher. According to this theory rate of interest is determined by the intersection of demand and supply of savings. It is called the real theory of interest in the sense that it explains the

determination of interest by analyzing the real factors like savings and investment. Therefore, classical economists maintained that interest is a price paid for the supply of savings.

Demand for Savings

Demand for savings comes from those who want to invest in business activities. Demand for investment is derived demand. Any factor of production is demanded for its productivity. The demand for the factor is high when there are higher expectations from it. Since, all the factors are not equally productive, so, capital demand will be high for more productive uses first and then gradually with the increase in its supply, will shift to less productive uses. Therefore, classical economists maintained that with the aid of capital facilities we turn out more goods per man-hour than when we produce with bare hands or with scant tools. Moreover, marginal productivity of the business goes on decreasing with more and more doses of investment of savings in his business venture. It is due to the operation of the law of diminishing returns.

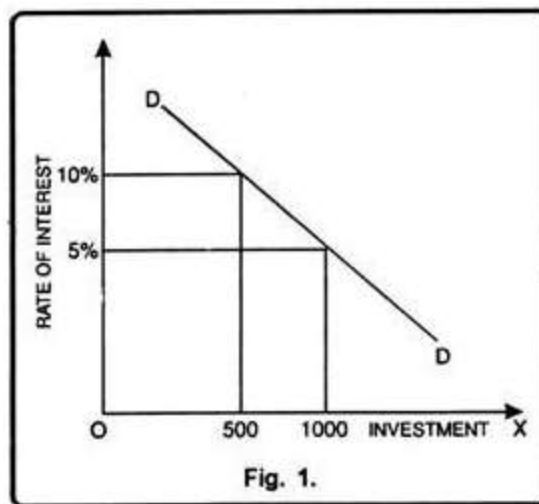
Now a very important question arises is that how much capital a person will demand because when a person borrows money he has to pay interest on it. The answer according to this theory is that demand for capital can be raised to a point where marginal productivity of capital becomes equal to the interest paid on it. Thus, if marginal productivity of capital is more than the interest paid, then it is beneficial to borrow money and vice-versa. Equilibrium will prevail at a point where marginal productivity of capital equals the rate of interest. This shows that there exists inverse relationship between demand for capital and the interest rate.

This fact can be made clear with the help:

Table 1

Rate of Interest	Investment (in crores)
10%	500
9%	600
8%	700
7%	800
6%	900
5%	1000

That the rate of interest and investment are inversely related to each other. As the rate of interest increases, the level of investment declines and vice-versa. As initially, the rate of interest is 10%, investment is Rs. 500 crores. When the rate of interest decreases to 8%, the level of investment increases to Rs. 700 crores from Rs. 500 crores. Further, as the rate of interest again falls to 5%, the level of investment increases to Rs. 1000 crores.



The depicts that there exists inverse relationship between the investment and the rate of interest. Initially, the rate of interest is 10%, the level of investment is Rs. 500 crores. Now the rate of interest falls to 5%. With this decrease in the interest rate, level of investment increases to Rs. 1000 crores. It indicates that more capital is demanded at a low interest rate and vice versa.

Supply of Savings

Supply of capital is the result of savings. It comes from those who have the excess of income over consumption. Thus, savings is the main source of capital which depends on the capacity to save, willingness to save, level of income and rate of interest etc. Capacity to save depends on the size of national income, size of personal income, size of family, price level and purchasing power of money etc. Willingness to save depends on the family affection,

further expectations etc. To a large extent, willingness to save is affected by the rate of interest. On a higher rate of interest people save more to earn the benefits of high rate of interest. On the other hand, at the low rate of interest, people save less. Thus, we may say that there is a direct relationship between the supply of savings and the rate of interest.

Table 2

Rate of Interest	Savings (in crores)
10%	1000
9%	800
8%	700
7%	600
6%	500
5%	400

The rate of interest and savings have a positive relationship. As the rate of interest increases, savings will also increase. On the other hand, a fall in rate of interest leads to a decrease in savings. When the rate of interest is 10%, the savings are of Rs. 1000 crores.

In the successive periods, as rate of interest falls from 10% to 5%, the total savings also decline. Suppose as the rate of interest falls to 5%, savings also decrease to Rs. 400 crores.

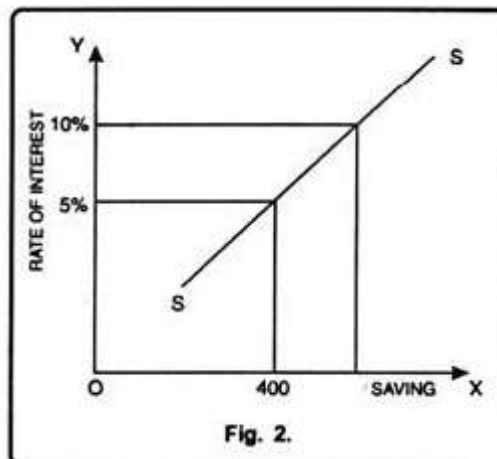


Fig. 2.

The savings have been represented on X-axis and interest rate on Y-axis. SS is the supply curve which moves upward from left to right. It shows that supply of savings is interest elastic. Higher the interest rate, more will be saved and vice-versa. With 5% rate of interest money savings are Rs. 400 crores. As the interest rate increases to 10% people are persuaded to save more and the money savings rise to Rs. 1000 crores. This signifies that there is a direct relationship between savings and the rate of interest.

Equilibrium Rate of Interest

According to classical theory, equilibrium interest rate is restored at a point where demand for and supply of capital are equal i.e.

Table 3

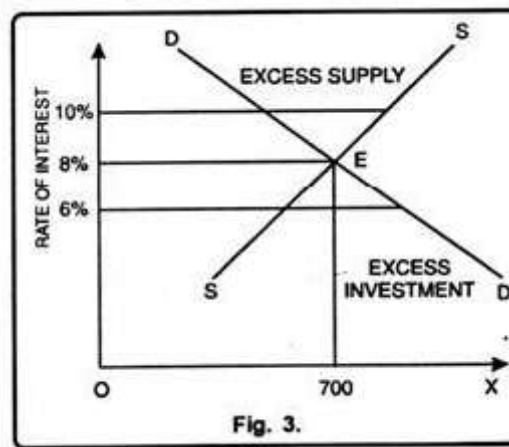
Rate of Interest	Investment	Savings
10%	500	1000
9%	600	800
8%	700	700
7%	800	600
6%	900	500
5%	1000	400

That equilibrium rate of interest will be determined at a point where demand for and supply of capital are equal. As is clear from the table that equilibrium interest rate 8% is determined because at this level demand for and the supply of capital are equal i.e. Rs. 700 crores.

Now, if the rate of interest increases to 10%, investment is Rs. 500 crores and savings are of Rs. 1000 crores i.e. savings exceed the investment. On the other hand, if the rate of interest falls to 5% investment is Rs. 1000 crores and savings are Rs. 400 crores.

The rate of interest is determined by the intersection of demand and supply curves. Equilibrium is restored at point E which determines rate of interest as 8% and demand and supply of capital as Rs. 700 crores. Now, if the rate of interest increases to 10% supply of savings exceeds the demand for capital i.e. supply is more than demand. This will lead to a fall in interest rate to the level of 8%.

On the other hand, when the interest rate falls to 6%, demand for savings exceeds the supply of savings which will push up the rate of interest to restore an equilibrium rate i.e. 8%. Therefore, rate of interest is in equilibrium only at a point where the demand equals the supply of capital.



Criticism

The classical theory of rate of interest has been criticized on the basis of the following shortcomings as discussed below:

1. Indeterminate Theory: Keynes has maintained that the classical theory is indeterminate in the sense that it fails to determine the interest rate. In this theory, interest is determined by the equality of demand and supply. But the position of savings varies with the income level. Thus, unless we know the income, interest rate cannot be determined.
2. Fixed Level of Income: Classical theory assumes that the level of income remains constant. But in actual practice income changes with a small change in investment. Thus, it is not correct to assume a fixed level of income.
3. Long Run: Classical theory determines the interest rate through the interaction of demand and supply of capital in the long run. Keynes pointed out that in the long run

we all are dead. Therefore, there was an urgent need of a theory which determines rate of interest in the short-run.

4. Full Employment: This theory assumes that there is full employment of resources in the economy. But, in reality, unemployment or less than full employment is a general situation. Full employment is only an abnormal case... Thus, this theory does not apply to the present world.
5. Savings and Investment: Classical economists assume that savings and investment are interring dependent. But actually investment changes, income also changes which leads to a change in savings. Thus, both are interdependent on each other.
6. Ignores Monetary Factors: Classical theory takes into consideration only the real factors for determining the rate of interest and ignores the monetary factors.

Profit

In contrast to a monopoly or oligopoly, it is impossible for a firm in perfect competition to earn economic profit in the long run, which is to say that a firm cannot make any more money than is necessary to cover its economic costs. In order not to misinterpret this zero-long-run-profits thesis, it must be remembered that the term 'profit' is also used in other ways. Neoclassical theory defines profit as what is left of revenue after all costs have been subtracted, including normal interest on capital plus the normal excess over it required to cover risk, and normal salary for managerial activity. Classical economists on the contrary defined profit as what is left after subtracting costs except interest and risk coverage; thus, if one leaves aside risk coverage for simplicity, the neoclassical zero-long-run-profit thesis would be re-expressed in classical parlance as profits coinciding with interest in the long period, i.e. the rate of profit tending to coincide with the rate of interest. Profits in the classical meaning do not tend to disappear in the long period but tend to normal profit. With this terminology, if a firm is earning abnormal profit in the short term, this will act as a trigger for other firms to enter the market. They will compete with the first firm, driving the market price down until all firms are earning normal

profit only. It is important to note that perfect competition is a sufficient condition for allocative and productive efficiency, but it is not a necessary condition. Laboratory experiments in which participants have significant price setting power and little or no information about their counterparts consistently produce efficient results given the proper trading institutions.

THEORIES OF PROFIT AND PROFIT POLICIES

Clark's Dynamic Theory of Profit And Its Criticism

The Dynamic Theory of Profits was introduced by J. B Clark. According to him profits come only in dynamic economy but not in static economy. Dynamic economy means the economy in which frequent changes will occur. In static economy there is no possibility of coming changes. In static economy there will not be any change both in demand and supply. So, profits cannot arise. According to Clark the following things exist in a static economy.

In such society the demand for goods can be estimated easily. There is no risk and uncertainty. Supply will be always equal to demand estimated. The price will be always equal to the cost of production, because there is perfect competition.

In static society the reward is just equal to the marginal productivity of the factors of production and so there cannot be profits. But we are not living in a static society. Ours is a dynamic world, where some changes are constantly taking place.

According to J. B. Clark, five main changes are constantly taking place in dynamic society. They are :

- 1) Changes in the number of human wants,
- 2) Changes in the methods of production,
- 3) Changes in the capital formation in the economy,
- 4) Changes in the method of organization of the business,
- 5) Changes in the size of population and incomes of the people.

According to Clark. profit arises in a dynamic society on account of these changes. These changes affect the demand for and

supply of commodities and thus lead to the emergency of profit. These are general dynamic changes. But sometimes dynamic changes may be introduced deliberately by the firms themselves. For example a firm may succeed in cutting down its cost by improving its production techniques and there by increasing its profit.

In short, it is the operation of dynamic changes which leads to the emergency of profit. According to Clark, profit belongs to economic dynamics but not economic statics.

Criticism of Dynamic Theory of Profits

This theory of profit has been criticized on the following grounds

1. Types of risks : According to Knight all types of dynamic changes do not yield profit. According to him two types of changes takes place in society. a) Foreseeable changes. b) Unforeseeable changes. According to him only unforeseeable changes brings profits.
2. Risk in undertaken by entrepreneur : According to Clark, risk is undertaken by capitalist but not entrepreneur, But it is not correct. If the capitalist is a risk taker, entrepreneur loose all his importance and become a salaried employee.
3. Determination of profit: The theory does not explain how the rate of profits can be determined.
4. Distinction between profits and wages : According to Prof. Taussig, dynamic theory has created an unnecessary and artificial wages of management.
5. Losses: According to Clark, whenever there is a change in the economy, it brings profit. But in reality it is not correct. Sometimes they may cause losses also.
6. Profit in the reward: This theory states that profits arises due to dynamic changes. It does not recognize that profits is the reward for entrepreneurs.

Schumpeter's Innovation Theory of Profit

Definition: The Innovation Theory of Profit was proposed by Joseph. A. Schumpeter, who believed that an entrepreneur can

earn economic profits by introducing successful innovations. In other words, innovation theory of profit posits that the main function of an entrepreneur is to introduce innovations and the profit in the form of reward is given for his performance. According to Schumpeter, innovation refers to any new policy that an entrepreneur undertakes to reduce the overall cost of production or increase the demand for his products.

Thus, innovation can be classified into two categories; The first category includes all those activities which reduce the overall cost of production such as the introduction of a new method or technique of production, the introduction of new machinery, innovative methods of organizing the industry, etc.

The second category of innovation includes all such activities which increase the demand for a product. Such as the introduction of a new commodity or new quality goods, the emergence or opening of a new market, finding new sources of raw material, a new variety or a design of the product, etc.

The innovation theory of profit posits that the entrepreneur gains profit if his innovation is successful either in reducing the overall cost of production or increasing the demand for his product. Often, the profits earned are for a shorter duration as the competitors imitate the innovation, thereby ceasing the innovation to be new or novice. Earlier, the entrepreneur was enjoying a monopoly position in the market as innovation was confined to himself and was earning larger profits. But after some time, with the others imitating the innovation, the profits started disappearing.

An entrepreneur can earn larger profits for a longer duration if the law allows him to patent his innovation. Such as a design of a product is patented to discourage others to imitate it. Over the time, the supply of factors remaining the same, the factor prices tend to rise as a result of which the cost of production also increases. On the other hand, with the firms adopting innovations the supply of good and services increases and their prices fall. Thus, on one hand the output per unit cost increases while on the other hand the per unit revenue decreases.

There is a point of time when the difference between the costs and receipts gets disappear. Thus, the profit in excess of the normal

profit disappears. This innovation process continues and also the profits continue to appear or disappear.

Profit: Meaning and Theories of Profit

We all are familiar with the term 'Profit'. It is quite a common-place word, but different people use it in different senses.

In Economics, however, the term has a precise meaning. Profit may be defined as the net income of a business after all the other costs – rent, wages and interest etc., have been deducted from the total income.

Profits are, therefore, uncertain and vary from person to person and from firm to firm. They may become zero, when costs are equal to income, and if the costs are higher, profits may actually be converted into loss.

Entrepreneur's Reward

Pure profit is the reward of entrepreneurial functions. It is what an entrepreneur gets purely as an entrepreneur. What he gets as a landlord, manager or capitalist is deducted from the total profits. Hence, Pure Profit is an amount which accrues to the entrepreneur for assuming the risk inseparable from business.

It is a reward for assuming the final responsibility, a responsibility which cannot be shifted to anybody else. A practical example will show how profits are calculated. Let us suppose that Mr. S. Chand starts a bookshop with a capital of Rs. 1, 00,000. We further assume that the premises belong to him and his nephew works for him without receiving any wages.

The above account shows the nature of profits. The meaning of gross profits too is clear. There are factors for which S. Chand does not pay. They belong to him, but ordinarily they could not have been obtained without payment. Therefore, these payments, though implicit, must be deducted to find out the net or pure profit of the business.

True Profit is thus a Residual Element

Profit is arrived at after the other three factors of production have received their remunerations out of national income. It may

become zero or even negative temporarily. But, in the long run, it must be positive, for otherwise the entrepreneur will give up his independent activity and take to service for wages.

Gross and Net Profits

We are now in a position to analyse gross profits.

They are the difference between total sale proceeds and total expenses over a year and include the following besides net profit:

- (a) Rent of the employer's land or premises: Had similar premises been taken on rent, the amount would have been added to costs. An equal figure should be deducted from gross profits to find out net profit.
- (b) Interest on entrepreneur's capital: The interest on borrowed capital is usually deducted before profits are worked out. Hence the interest which the owner's own capital would have earned elsewhere should be taken out before we can determine net profits.
- (c) Wages of management: The entrepreneur may be himself providing the services of management. If he had been employed elsewhere he would have earned some wages. An equal amount has; therefore, to be deducted before net profits can be discovered.
- (d) Maintenance charges: It is but reasonable that capital should be maintained intact. Worn-out pieces should be replaced at the proper time. To do this, it is necessary to maintain a depreciation fund. All expenses for this purpose should be deducted out of gross profits. If this is not done, profits will appear to be large for a few years but one day the business will fail, because no funds will be available to replace fixed assets like machinery.
- (e) Net Profits: If we deduct from gross profits the above items, we shall get pure or net profits. The entrepreneur is entitled to the following different kinds of payments which form a part of his net profits.

Reward for risk-taking

Every business faces some risk of loss. But the risk of loss

from market-fluctuations has to be borne by the entrepreneur himself, and he will shoulder it only when he has hopes to be paid for it.

Reward due to a monopolistic position

A particular entrepreneur may earn extra income due to his control in the market over the entire supply of the commodity he produces.

Rent Theory of Profit

The Rent Theory of Profit was propounded by an American economist F.A. Walker. He was the first to introduce a distinction between a capitalist and an entrepreneur into English economic theory. An entrepreneur need not be a capitalist. He is a person who may undertake a business without using any of his own capital.

Rent of Ability

Walker regards i profit as rent of ability. Just as there are different grades of land, there are different grades of entrepreneurs. The least efficient entrepreneur, who must remain in the field of production to meet the current demand, just recovers his cost of production and nothing besides.

Above him are entrepreneurs of superior ability. Just as rent arises because of the differential advantage enjoyed by superior land over the marginal land, similarly profit also is the reward for differential ability of the entrepreneur over the marginal entrepreneur or the no-profit entrepreneur.

Profit is thus like rent and, like rent it does not enter into price. Wages of management are not profit. The marginal employer only earns the wages of management, and no more. With a slight unfavorable turn of prices or costs, he would prefer to work as an employee rather than as an employer. Wages of management thus must be paid to keep up the given supply of entrepreneurs. Such wages thus enter into price.

Bibliography

- Bardsley, N., Cubitt: *Assessing Experimental Economics*, Princeton, Princeton University Press, 2009.
- Bazilevskii, Y. Y.: *The Theory of Financial Management*, Pergamon Press, Macmillan Co., New York, 1963.
- Borcherding, K., : *Contemporary Issues in Business Decision Making*, Amsterdam, North-Holland, 1990.
- Desmond, Glenn, and John A, Marcell, *Handbook of Small Business Valuation Formulas and Rules of Thumb*, Valuation Press, 1993.
- Dimson, Elroy: *Stock Market Anomalies*, Cambridge University Press, 1988.
- Fruhan, W. E.: *Financial Strategy, Studies in the Creation, Transfer and Destruction of Shareholder Value*. Homewood, 1979.
- Gowdy, J., *Coevolution Economics: The Economy, Society and the Environment*, Kluwer, Boston, 1994.
- Hausman, D. M.: *The Inexact and Separate Science of Economics*, Cambridge, Cambridge University Press, 1992.
- John A, Marcell, *Handbook of Small Business Valuation Formulas and Rules of Thumb*, Valuation Press, 1993.
- Kasper, Larry J.: *Business Valuations, Advanced Topics*, Westport, CT, Quorum Books, 1997.
- Knapp, C. L.: *Commercial Damages, A Guide to Remedies in Business Litigation*, Matthew Bender, 1993.
- Larry J.: *Business Valuations, Advanced Topics*, Westport, CT, Quorum Books, 1997.
- Lindblom, C. E.: *A Strategy of Decision: Policy Evaluation as a Social Process*, New York, The Free Press, 1970.
- Loomes, G.: *Current Issues in Microeconomics*, New York: St. Martin's Press, 1989.
- Marris, R. L., and Wood, A.: *The Corporate Economy*, London: Macmillan, 1971.
- Meyer, M. W.: *Theory of Financial Organizational Structure*, Indianapolis, 1977.

- Pennings, J. M.: *Decision Making, An Organizational Behaviour Approach*, New York, Markus Wiener Publishing, 1983.
- Pesendorfer, W.: *The Foundations of Positive and Normative Economics*, New York, Oxford University Press 2008.
- Rappaport, A.: *Creating Shareholder Value, The New Standard for Business Performance*, New York, Free Press, 1986.
- Robert Smith: *Modern Labor Economics*, Harper Collins, 1994.
- Roth, A. E.: *The Handbook of Experimental Economics*, Princeton, Princeton University Press, 1995.
- Santos, A.C.: *The Social Epistemology of Experimental Economics*, London, Routledge 2009.
- Schumpeter, J. A., *The Theory of Economic Development*, Harvard University Press, Cambridge, Massachusetts, 1934.
- Sinha, R.K.: *Economic Development since Independence*, New Delhi, Deep and Deep, 1988.
- Taylor, F. W.: *The Principles of Business and Finance*, New York, 1917.
- Throsby, David, *Economics and Culture*, Cambridge University Press, 2001.
- Tietenberg, Tom, *Environmental Economics and Policy*, New York, HarperCollins, 1994
- Tinbergen, Jan.: *On the Theory of Economic Policy*; Amsterdam, North-Holland 1952.
- Ullman, D. G.: *Making Robust Decisions*, Trafford, 2006.
- Weber, M.: *The Theory of Social and Economic Organizations*, New York, Oxford University Press, 1947.
- Weibull, J. W.: *Advances in Understanding Strategic Behaviour*, New York, Palgrave, 2004.
- Woodward, J.: *The Oxford Handbook of Philosophy of Economics*, New York, Oxford University Press, 2009.
- Yegge, Wilbur M., *A Basic Guide for Valuing a Company*, New York, Wiley, 1996.
- Zipp, Alan S., *Business Valuation Methods* New York, American Institute of Certified Public Accountants, 1990.

Index

A

Action Plan, 94.

B

Basic Problems of Economy, 1.
Business Process Outsourcing,
24, 25.

C

Change in Demand, 80, 83, 84,
85, 101, 108, 109, 115,
272.
Characteristics of Monopoly,
235.
Classification of Cost, 188.
Concept of Production
Functions, 131.
Constant Returns to Scale, 172,
173, 176.
Cost Analysis, 182.

D

Demand Analysis, 72.
Demand Estimation, 85.
Demand Forecasting, 88, 95,
96, 97.
Determinants of Demand, 28,
73, 76, 84, 89.
Diminishing Returns to Scale,
172, 173, 175.

Diseconomies of Scale, 159, 161,
162, 166.

E

Economic Structure, 17, 19.
Elasticity of Demand, 99, 104,
108, 109, 110, 112, 115,
119, 120, 121, 122.
Elements of Cost, 184, 187, 208.
Empirical Estimation, 85.
Expenditure Method, 31, 36,
40, 44, 56, 57, 63, 64.
Expenses of Selling Costs, 252.

F

Factor Pricing, 270, 271, 275.

G

Government Statistics, 19.

I

Income Statement, 91.
Increasing Returns to Scale, 170,
174.

L

Law of Demand, 19, 74, 81,
82, 83, 91, 95, 107, 112.

M

Managerial Decision Making, 4.
 Market Price, 47, 48, 58, 273.
 Measuring National Income, 31,
 32, 40, 41, 42, 46, 49,
 66.
 Monopolistic Competition, 214,
 243, 244, 245, 249, 250,
 252, 254, 259, 266.

N

National Income, 29, 30, 31,
 32, 38, 40, 48, 52, 56,
 68, 69, 70, 71, 72.

O

Organizational Situations, 4.
 Output Determination, 217,
 238, 259.

P

Perfect Competition, 213, 214,
 216, 223, 232, 235, 259.

Price Determination, 112, 213,
 237, 245.

Production Function, 124, 127,
 133, 136, 232, 233, 234.

S

Small Businesses, 2.
 Social Responsibility, 6.
 Standard Cost Accounting, 206.
 Supply Schedule, 89.

T

Techniques of Demand
 Forecasting, 97.

V

Variable Proportions, 138, 147,
 148, 151.
 Various Economic Methods, 10.

W

Wage Theory, 281.

□□□