PAGES MISSING WITHIN THE BOOK ONLY

(113 TO 148 & 341 TO 364)

UNIVERSAL
LIBRARY

OU_152798

UNIVERSAL
LIBRARY

Fundamentals of Economics

The quality of the materials used in the manufacture of this book is governed by continued postwar shortages.

Fundamentals of Economics

by MYRON H. UMBREIT, PH.D.

Director of the Chicago Division of the School of Commerce, Northwestern University

ELGIN F. HUNT, M.A.

Coordinator of the Social Sciences, Chicago City Junior College; Lecturer in Economics Northwestern University

CHARLES V. KINTER, PH.D.

Business Economist
Lecturer in Economics, Northwestern University

FIRST EDITION

McGRAW-HILL BOOK COMPANY, INC.

NEW YORK TORONTO LONDON • 1948

FUNDAMENTALS OF ECONOMICS

Copyright, 1948, by the McGraw-Hill Book Company, Inc. Printed in the United States of America. All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

Preface

"Fundamentals of Economics" was developed in a laboratory. That laboratory was the classes in economics on the Chicago campus of Northwestern University, where the book, in its earlier forms, has been used for several years. The present edition was preceded by three mimeographed or planographed editions, and each new edition represented an extensive revision. Each time large portions of the book were completely rewritten. These successive revisions were based on discussion with students, the criticism of colleagues, and above all, on our own classroom experience.

"Fundamentals of Economics" is intended to introduce college students to the study of economic principles. Therefore, the discussion has been limited as far as possible to those basic facts, concepts, relationships, and institutional arrangements which are most essential to an understanding of our economy. Every effort has been exerted to make the discussion clear, concrete, and to the point. To this end illustrations and charts have been used freely whenever it was felt that they would help the reader.

On the basis of their own experience the authors believe that "Fundamentals of Economics" can be used to advantage in any first course in economic principles. Though the book is modern in its viewpoint and its theories, the general organization is along traditional lines. The authors have not been interested in innovation for its own sake. They believe that new developments in economics can be best understood if they are brought into relationship with those developments which came earlier. Their primary purpose has been to write a text that is (1) clear and readable, (2) capable of interesting the student, (3) sound in theory, and (4) thoroughly teachable. To aid both student and teacher each chapter is provided with a summary, and with questions and exercises. References are also given for the benefit of students who wish to do additional reading.

The authors wish to express their deep gratitude to all those friends and associates who read the manuscript in whole or in part and who made valuable suggestions. However, responsibility for the ideas expressed and for any errors that may have crept in is wholly the authors'.

THE AUTHORS

CHICAGO, ILL.
February, 1948

Contents

PREFACE v

1. *Introduction* 1

Role of the consumer, Social implications of business activity, Necessity for profits, Conditions essential for carrying on business, Business problems and economic problems.

2. *The Foundations of Economic Activity* . . . 12

The Nature of Man—The Nature of Man's Physical Environment—The Nature of Organized Society; *Relation of the individual to society;* Three Characteristics of Modern Economic Society; *Voluntary cooperation, Competition, Monopoly;* Competition and Monopoly Compared; *Advantages of competition, Disadvantages of competition, Advantages of monopoly, Disadvantages of monopoly;* Summary.

3. *Fundamental Concepts* . 28

Good, Utility, Free goods, Economic goods, Services, Wealth, Producers' and consumers' goods, Production, Factors of production, Money, Income, Types of income, Property, Public versus private property, Value, utility, and price; Summary.

4. *The Organization of Production* 35

Nature of production, Form utility, Place utility, Time utility, Possession utility, Service utility; The Nature and Functions of Management; *The entrepreneur or businessman;* Size of the Business Enterprise; *Determining the size of an enterprise, Advantages of large-scale production, Limitations to large-scale operation;* Forms of Business Organization; *The single proprietorship, The partnership, The corporation, Disadvantages of corporate organization;* Summary.

5. *Money* 50

The nature of money, Functions of money; Monetary Standards; *Gold standard, Bimetallism, Paper standards;* Bank Money; *Bank notes, Bank deposits;* Money in the United States—Summary.

6. *The Commercial Banking System* . 64

Commercial banks, Bank reserves; Our Banking System Before and After 1913—Federal Reserve System; *Reserve banks and the member banks, Federal Reserve bank reserves, Federal Reserve notes, How bank deposits are created;* Summary.

7. *The Purchasing Power of Money* 80

The value of money; How the Value of Money Is Determined; *Equation of exchange, The quantity theory;* Measuring Changes in the Price Level; *Index numbers, Method of weighted aggregates, Purchasing power of money;* Effects of Changes in the Value of Money; *General price level and individual prices, Business activity and the rate of spending;* Summary.

8. *Income, Expenditure, and Employment* . . . 98

Relation of Income to Expenditure—Consumption, Saving, and Investment—Propensity to Consume and Rate of Saving—Variations in Prices, Production, and Employment—Multiplier Effect—Effect of Speculation on Expenditure, Income, and Employment—Level of Employment—Inherent Instability of Our Economy—Basic Conditions of Full Employment—Summary.

9. *The Price-making Process: Demand* . 114

Nature of demand, Diminishing utility and demand, Marginal price offer, Demand schedule illustrated, Demand curve, Changes in the demand schedule, Elasticity of demand; Summary.

10. *The Price-making Process: Supply and the Equilibrium Price* 131

Meaning of supply, Supply schedule, Supply schedule chan[illegible] Equilibrium of Demand and Supply; Changes in demand and supply schedules, Perfect competition, Time factor in price determination; Summary.

11. *Cost of Production* 150

Cost of production and price, Meaning of cost of production, Opportunity costs, Cost, prices, and incomes, Short-run relationship of cost to price, Relation of costs to prices in the intermediate period, Fixed and variable costs, Marginal cost, Fixed and variable costs in relation to output, average cost, and profits; Summary.

12. *Proportionality* 166

Cost curve of the individual firm, Principle of diminishing returns, Diminishing returns and costs of production, Principle of proportionality or best combination, Principle of variable proportions; Summary.

13. *Price in the Intermediate Period* 183

Demand schedules in the intermediate period, Supply in the intermediate period, Demand for the product of the individual producer, Equilibrium price in the intermediate period, Differences in average costs, Effect of price changes on the individual firm, Nature of the tendency of marginal cost to equal price; Summary.

14. *Price in the Long Run* 192

Nature of the long-run period, Prices in the long-run period, Tendency of average costs to become equal, Long-run cost trends in different industries, Increasing-cost industries, Constant-cost industries, Decreasing-cost industries, Relation of long-run cost trends to price; Summary.

15. *Price under Dynamic Conditions* 206

Price determination under dynamic conditions, Business adjustments in a dynamic market, Representative firm, Representative firm under dynamic conditions, Representative firms in actual competitive markets, Bulk-line cost; Summary.

16. *Price under Monopoly* 221

Fixed stock, Reproducible stock, Average revenue, Marginal revenue, Monopoly price in the short run, Monopoly price in the intermediate period, Monopoly price in the long run, Monopoly price under constant cost, Monopoly price under increasing cost, Monopoly price under decreasing cost, Class price; Summary.

17. *Imperfect Competition* 244

Partial Monopoly; Duopoly, Oligopoly, Partial monopoly in practice, Examples of duopoly, Examples of oligopoly, Monopolistic competition, How product differentiation can be achieved, Why product differentiation is sought; Summary.

18. *Interdependent Prices* 260

Prices under joint demand, Prices of substitutes, Partial, complete, and potential substitution, Prices of joint-cost products, Prices of inseparable-cost goods, Charging what the traffic will bear; Summary.

19. *Economics of International Trade* 273

Differences between domestic and international trade, Advantages of world trade, Attitude of nations toward world trade, The stake of the United States in world trade, Agencies designed to foster world trade, Basis for trade between nations, Principle of comparative advantage, Modifications of the principle of comparative advantage, Mechanics of foreign trade, Bills of exchange, Gold points, Buyers and sellers of bills of exchange, Indirect exchange, Favorable and unfavorable balance of trade, Visible and invisible items in trade, Why trade between countries must balance in the long run, Foreign trade with nations off the gold standard, Trade restrictions, Valid arguments in favor of imposing trade barriers, Invalid arguments in favor of imposing trade barriers; Summary.

20. *Distribution* 302

Meaning of distribution, Functional distribution and personal distribution, Prices of the factors of production, Productivity as the source of the demand for a factor, Marginal productivity, Marginal productivity schedule, Market demand for a factor of production, Market supply of a factor of production, Determination of the price or share of a factor, Prices of unlike units of a factor of production, Simultaneous determination of wages, interest, rent, and profits, Some implications of distribution theory, How the national income of the United States is divided; Summary.

21. *Rent* 320

Rent defined, Kinds of rent, Problem of explaining rent, Demand for land, Supply of land, Determination of the rental price, Rent on unlike units of land, Rents on different units of land under extensive cultivation, Marginal land, Differences in rents when cultivation is intensive as well as extensive, Formula for finding the rent of a unit of land, Relation of average cost to marginal cost and rent, Rents on residential and commercial land, Effect of rent on prices, Capital value of land; Summary.

22. *Wages* 341

Fundamental principles of wage determination, Immobility of labor, Noncompeting groups, Psychic income, Determination of wages under competition, Nature of the demand for a given kind of labor, Supply of a particular kind of labor, Law of substitution, Limitations on competition in real labor markets, High wages in the United States, Subsistence theory of wages, Wages-fund theory, Bargaining theory of

wages, Exploitation theory of wages, Summary of wage theories, Labor unions and wages, Real wages versus dollar wages, Wage rates versus annual income; Summary.

23. *Interest* 365

Interest defined, Contractual interest versus the pure rate of interest, Interest rate, Demand for capital, Supply of capital, Foresight, Determination of the interest rate illustrated, Interest rate and the quantity of money, When competition is not perfect; Summary.

24. *Profits* 382

Role of the entrepreneur in a free society, Other payments that may be included in profits, Relationship between risk and profits, Distinction between the use of the term "profit" by the accountant and by the economist, Profits earned in American industry, Criticism of profits, Entrepreneur in a corporation; Summary.

25. *Taxation* 395

Why has the cost of government risen?, Sources of government revenue, Characteristics of a well-designed tax system, Possible principles of apportioning costs of government among taxpayers, Formulation of tax rates, How high can tax rates go?, Question of incentives, Taxes for nonfiscal purposes, Kinds of taxes, Fiscal systems of national, state, and local governments, Shifting and incidence of taxes, Federal debt; Summary.

26. *Labor Organization* . . . 415

Development of labor organization in the United States, Reasons for the development and growth of labor unions, Structure of organized labor, Collective bargaining, Weapons of organized labor, The company union versus the independent union, The closed versus the open shop, Anti-union activities of employers, Norris–La Guardia Act, Labor under the New Deal, Labor in wartime and in the postwar period, Taft-Hartley Act, Economic effects of the growth of organized labor; Summary.

27. *Role of the Consumer* 432

The consumer and the price system, Direct consumption and indirect consumption, Foresight, Consumer choices, Factors affecting the consumer's freedom of choice, Authoritative control, The choice between spending and saving, Propensity to consume; Summary.

INDEX 451

1. Introduction

Economics is a study of the social aspects of business. It deals with the activities of businessmen and with the principles and practices involved in producing, exchanging, and distributing economic goods. It is important, at the outset, that the scope of the study be clearly defined. Economic theory, in its broadest sense, includes all the activities in which individuals engage in order to satisfy their wants for scarce goods and services. Such a study, however, encompasses so broad a scope that it would be impossible to include in one book all of the material necessary for an understanding of economic life. Our study, therefore, will be limited for the most part to the fundamental activities in which individuals engage for the purpose of making a living. Because of the important role of the businessman in a free-enterprise system, we shall devote especial attention to his endeavors to earn profits.

Since the businessman occupies such a prominent place in our economy, before we begin this study of economic activities and the forces that control them it is important to know just who is meant by the businessman or enterpriser; or, to use a term often found in economic literature, the "entrepreneur." For our study, the term "businessman" must be defined in a narrow sense. (*It refers only to those who promote business enterprises, who dictate policy, and who exercise ultimate control.*) They are the real businessmen; all others are merely employees who carry out directions.

The businessman, it is said, is the final coordinator of the factors employed. He takes the ultimate risks and makes the final decisions; in other words, he "dictates business policy." [1] It must not be assumed, however, that the businessman or enterpriser is the only one who makes decisions. Managers, as distinct from the enterpriser, also make decisions. The consumer is constantly making decisions that the businessman cannot

[1] GARVER and HANSEN, "Principles of Economics," Ginn and Company, Boston, 1937, p. 43.

ignore. Investors make decisions. The government passes laws that affect the decisions of businessmen. Workers have a voice in wages, hours of labor, and the like. Customs and habits, ways of doing things, and precepts of right or wrong action all limit the decisions of the businessman. Finally, the actions of competitors vitally affect the businessman's freedom to dictate business policy.

Moreover, the businessman is not the only person who takes risks. Labor assumes risks of unemployment, accident, and technological change. Investors assume risks, and hired managers and administrators assume risks. The businessman, however, assumes a special kind of risk, since by taking the responsibility for a venture he stands generally to lose both his reputation and his capital if the venture fails. We see, therefore, that, while the term "businessman" as used in this study refers to a very limited group of individuals, yet this group plays a part that is of vital concern to all society.

Though our principal concern in this book is with the profit-making activities of businessmen, we cannot entirely overlook either the consumer or the effects of business upon society as a whole. We shall, therefore, in this introductory chapter, briefly survey both the role of the consumer and the effect of business activity upon social welfare.

Role of the consumer. Important as is the businessman in our economic pattern, his influence and power would disappear if it were not for the consumer. Without consumers, no business can exist. In order to illustrate the significance of the consumer, let us take a simple example to show the part he plays in the economic process.

Mrs. X, a housewife, enters a corner grocery store to purchase a loaf of bread. This action in itself is a very simple one, but to make it possible for Mrs. X to purchase bread a whole chain of events must previously have occurred. The farmer must have cleared and cultivated land, planted and harvested wheat, and sold it to the grain dealer. He in turn must have sold it to the miller, who ground it into flour which was bought by the baker. The baker prepared and sold the bread to the grocer, who now has it ready for Mrs. X to purchase.

We should note that every activity prior to the actual purchase of bread by Mrs. X was made in anticipation of that purchase. We might even have gone back further. Before the farmer cultivated his land, farm implements must have been manufactured, and before that, iron ore and coal must have been mined. These and many other activities that might

have been mentioned were all engaged in because of the demands of consumers for bread. Because consumption is the final goal of production, the consumer by purchasing or refusing to purchase determines the kinds and quantities of goods that are to be produced.

The consumer is also sensitive to price changes. For example, when Mrs. X entered the store and asked for a loaf of bread she expected to pay a certain price for it, let us say 10 cents. Suppose the grocer has previously sold bread for that price, but today he has raised the price to 12 cents. What explanation will he give to Mrs. X and what will be her reaction? Will she purchase the bread, or will she purchase a substitute product, such as potatoes? It is clear that the price of a consumer's good is vital in determining the quantity that will be purchased. It is the consumer, therefore, who determines how much of a good can be sold at various prices, and it is also the consumer who determines the kinds of goods that can be sold.

The decision of a consumer as to what and how much to purchase depends largely on income. Suppose Mrs. X's husband is employed and is receiving a wage of $200 per month. On the basis of this wage she has decided on a pattern of purchases at prevailing prices. However, if Mr. X's wages are cut or prices increase, she must change this pattern; since she can spend fewer dollars, or must pay higher prices, she must resort to one of the following two alternatives: She may reduce her purchases of all goods to a greater or less degree; or she may reduce her purchases of some goods in order to buy more of others. For example, she may buy more of relatively cheap foods in order to substitute these for expensive foods that she can no longer afford.

Since the housewife has a limited income and since in spending that income she will seek to secure the greatest amount of desirable goods, it is clear that the businessman is very largely dependent on the pattern of consumer expenditure. Whether he is planning plant expansion or production schedules or a sales campaign, he must always keep in mind not only the consumer's needs and desires, but also the consumer's income.

Social implications of business activity. In countries like the United States private business is permitted, and the rights of the businessman are protected by law. The theory behind this is that activities of businessmen are, in general, beneficial to the community. In return, however, for his rights and privileges the businessman has social responsibilities. The

community looks to him to produce the goods that consumers desire. It also looks to him to reduce costs, to increase the quality and quantity of his products, and to share the benefits of increased production with the workers, the owners of capital, and the consumers. Unless business as a whole benefits society, businessmen are not entitled to the incomes that they receive, for a businessman should receive a reward only if he is rendering a service to society.

At times, however, a few businessmen enhance their incomes because they have discovered ways and means to exploit others. Sometimes they use monopoly power to extract excessive prices from the consumer. Sometimes they are able to hire labor at substandard wages. Or again, by financial manipulation they are able to return to the investor an amount below the actual earnings on his capital. These cases, however, are the exception rather than the rule, for, by and large, businessmen have lived up to their social responsibilities. They have promoted the general welfare by making possible an ever greater production of goods at ever lower prices.

Although the record of business in this respect has been highly commendable, there is still room for improvement. In consequence it has been suggested that, over the years, certain tests should be applied to every business concern to determine its contribution to the social welfare. It has been suggested that the following questions must be answered in the affirmative if a business is to be classed as contributing to the general welfare. First, has the business increased wages? Second, has it increased employment? Third, has it improved the quality of the product? Finally, has it lowered prices to the consumer? If a business has achieved some of these goals, without doing so at the expense of others, then clearly it is contributing to the general welfare. If it has not achieved any of these goals, it is possible that it is following socially undesirable policies which it should be required to modify.

An illustration may aid in making this point clear. It has been stated that, in the long run, it is the responsibility of businessmen to increase consumer incomes. This can be achieved by lowering prices and by increasing wages and dividends. If either prices are lowered or wages or dividends increased, some gain in the national welfare will result To demonstrate the possibilities of such gains we shall cite an example showing what was accomplished in one business.

In 1932 a certain producer had total sales per employee of only $5,500

per year. He could pay an average annual wage of only $1,300 and a dividend of only $2 per share of stock. The volume of his business was so small that he could employ only 200 men. Between 1932 and 1943, however, he was able to expand his business and achieve great operating economies. As a result, he was able to increase annual productivity from $5,500 to $28,000 per worker; moreover, this increase was achieved even though over the 11-year period he reduced the price of his product from $550 to $190 per unit. Because of the increased productivity per worker he was able to increase annual wages from an average of $1,300 to an average of $5,400 and to increase dividends from $2 to $6 per share. In other words, by increasing the output per worker he was able to increase wages and dividends while at the same time he reduced prices. Thus all groups benefited, and public welfare was promoted.

Of course, some businessmen have more ability than others, and some are more favored by circumstances. Nevertheless, experience has demonstrated that industry by and large can make tremendous strides in bringing more and better goods to the market at lower costs per unit.

Necessity for profits. Although businessmen must recognize their social responsibilities, this must not be interpreted to mean that they should disregard profits. The businessman who does not have a keen sense of money values will find himself taking heavy losses, and very soon he will lose his business.

The word "profits" is a much abused term. Often it is assumed that profits represent an income that has not been earned. Businessmen, however, use the term merely to indicate any income that they receive after all costs have been paid. In some cases such income may be great, in others very small. Large profits may or may not be a well-deserved return for services rendered. In general, however, business profits represent just rewards. Every business must assume certain risks, such as those arising from price changes. Business losses are often heavy, and the mortality rate among businesses is high. In consequence, returns must be substantial to those who are successful, for otherwise it would not pay to take the risks. We must remember that profits do not arise because people take risks; rather, people take risks in the hope of securing profits. Although it is true that at times profits accrue to businessmen because of unforeseen and lucky circumstances, like a sudden rise in prices, yet in the main profits accrue only to those who exercise skillful management and shrewd judgment.

Later in the discussion the question of reasonable and unreasonable profits will be considered. At this point we shall assume that most profits are justifiable rewards which businessmen receive in return for anticipating the wants of consumers, producing goods or services to supply those wants, and taking the risks that are necessarily involved. It is the search for profits that induces the businessman not only to expand production but also to increase the employment of labor and capital, and to make possible higher wages and dividends. Often, also, the search for profits induces the businessman to reduce prices or to improve the quality of his product. Larger sales and larger production may mean lower costs per unit; and even if the profit margin per unit falls, this may be much more than offset by the sale of a greater number of units.

Conditions essential for carrying on business. Under a free-enterprise economy, as we have seen, society looks to the businessman to provide for the wants and needs of consumers. It also looks to him to furnish goods in constantly increasing quantities and at ever lower prices. In the long run, if the businessman is to serve the community well, he must be permitted to make reasonable profits. However, merely to permit him to make profits is not enough. He must be permitted to make them under such conditions that his activities tend to benefit rather than injure the community.

The first basic condition for efficient business is reasonable stability. Businessmen must make plans in the present in anticipation of future needs. Such planning is almost impossible if the conditions surrounding economic activity are constantly changing in unpredictable ways. If goods are to be produced next year or the year after, decisions must be made now. Capital must be borrowed, plants built, and labor trained. If the necessary capital is to be obtained, investors must have some reason for believing that a project is likely to be a success, and this means that estimates must be made in the present of future costs, future volume of sales, and future selling prices. The greater the degree of stability under which he operates, the greater is the confidence of the businessman in his plans. Fear and uncertainty tend to bring business stagnation; confidence promotes energy and enterprise. Perfect stability is not of course possible, perhaps not even desirable. Some uncertainties always remain, and businessmen must always accept a certain amount of risk.

Although the businessman, through the exercise of foresight and intelligent planning, may wield considerable influence in securing and

maintaining a reasonable degree of stability, other influences may be present that may nullify his efforts in this respect. Political action may upset his plans, especially if the government follows an erratic and unpredictable program. Often, in its attempt to improve conditions, the government passes laws without understanding their effect upon the business pattern. If these laws are passed without adequate warning, or if they are imposed without good reason, the businessman finds himself in a difficult position. When changes are necessary in the interest of the general welfare, every effort should be made to avoid unduly disrupting the pattern of business activities. This can be accomplished if definite policies are established and if every effort is made by the government to adhere to such policies.

A second basic condition, which usually tends to foster business enterprise and guide it into socially useful channels, is a free and competitive market. The actual markets in which businessmen produce and sell their products vary greatly in type. Some markets are truly competitive; others are dominated by monopolists. In some markets buyers and sellers are free to bargain; in still others prices and the conditions of sale are subject to government regulation. In our system of free enterprise most kinds of business will provide the greatest social benefits if they are required to produce and sell their products in free, competitive markets. In such markets production and prices are flexible, for anyone can produce whatever he desires and offer it for sale. In markets in which free competition prevails, only those producers who can operate efficiently and sell at a relatively low price will continue in business; all others will fail. Competition forces the businessman to conserve man power as well as raw materials. It also induces him to expand production and sales. Under these conditions employment, money incomes, and real purchasing power all tend to be maximized, and as a result, the general welfare is promoted.

Although most kinds of business serve the public best when operating under conditions of competition, there are some exceptions, notably the public utilities. In these industries very large plants and extensive equipment are generally necessary if they are to operate efficiently; and since they can operate only in a limited market, if competition were encouraged, duplication of plant and equipment would lead to extensive waste and very high prices to consumers. Moreover, since the utilities require rights of way, duplication of equipment would needlessly clutter

our streets and our rural landscapes. It is now generally recognized that the public utilities should be operated as monopolies, but that, in order to protect the public, they should be under government supervision.

In those industries in which monopoly is advantageous to the public, government policy should not only permit but encourage it. On the other hand, in that larger group of industries where competition best serves the public interest, government policies should be designed to protect competition and put obstacles in the way of monopoly.

A third basic condition for the effective functioning of business is the existence of an efficient and mobile labor force. Labor is the active factor of production. In order to take advantage of its greatest opportunities labor must be not only well trained but also adaptable to changing circumstances. To be most productive and receive the highest possible wages, sometimes even to be employed at all, a worker must be able to shift from one industry to another as new products replace old ones. Also, he must be able to shift from one job to another as methods of work change; and he must be able and willing to move from one part of the country to another as industries change their location.

This problem of the mobility of labor is one of the greatest under our present economic system. Changing jobs is often difficult, especially for trained workers whose skills are no longer in demand because of changed industrial needs. Inducing workers to move from one place to another in order to better themselves is also difficult. Labor is a human factor, and a worker is often bound to his locality by strong ties of sentiment and by family and friends. Nevertheless, a reasonably mobile labor force is very necessary if we wish to increase production and raise standards of living.

The basic requirements, then, for efficient and socially useful business activity are these three: (1) reasonably stable conditions under which to carry on business; (2) competitive markets for most industries, with monopoly under government regulation encouraged in certain cases where it best serves the public interest; (3) an efficient and reasonably mobile labor force. Where these conditions are approximated, most business activities should bring benefits to both the businessman and the community.

Business problems and economic problems. Every businessman faces problems in the operation of his business. The very complexity of business relationships makes this inevitable. Some of these problems result from the businessman's own ignorance, mistakes, and shortsightedness; others

grow out of the difficulty of reconciling short-run and long-run objectives. Some business problems concern the internal organization of an enterprise as an efficient productive unit; others, again, concern the relation of the individual enterprise to the rest of the economic system.

To a large extent the problems that the businessman faces in operating his own enterprise are created by maladjustments in the economic system as a whole. These maladjustments are not business problems in the narrow sense, because there is very little that the individual businessman can do about them. Because they affect the entire economy, they are properly called *economic problems*. As a rule, economic problems can be solved only by group action, and sometimes the only agency in a position to take such action is the government.

Much of our study in economics will be devoted to basic problems that must be solved if we are to make progress. As we progress beyond the elementary level, we shall find that the problems are many and cover a wide field. They relate to such matters as the business cycle, inequality of wealth and income, inefficiency in production and marketing, foreign trade, social security, and economic planning. To the extent that any one vital problem is solved, to that extent business can make more progress toward achieving its goals. In our study we shall frankly recognize that serious economic problems exist and that there are good reasons for their existence. However, we shall also assume that they can be solved, and that to solve them it will not be necessary to change the form of our government or to abandon our system of free enterprise in favor of some other system.

We must remember, however, that there is seldom any simple solution for an economic problem. Since most economic problems are complex, there is danger that if we do not really understand them we may apply the wrong remedies. This will merely create new problems and may create even greater difficulties than existed before the "remedy" was applied.

Because economic problems are generally so complex, it may be that no simple direct solution can be discovered. Instead of seeking to attack a problem directly it is often necessary to correct conditions that bear upon it only indirectly. A roundabout approach is often more effective than a direct assault.

The solution of economic problems may be compared to the actions of the skipper of a sailboat who wishes to reach port. He is entirely

dependent upon wind power, and yet the port that he wishes to reach is in the very direction from which the wind is blowing. To the landsman his situation might seem hopeless, for it is obvious that he cannot sail directly into the wind. However, those who understand sailing know that by following a course sometimes to the right and sometimes to the left he will eventually reach his goal. Just as the straightest course is not the way for the skipper to reach port, so the simplest and most obvious procedure is not always the way to achieve economic goals.

In order to analyze an economic problem and find the solution for it we must first know something of the real world in which business is carried on. This, however, is not enough. We must also have a thorough grasp of basic economic principles. In our study, therefore, we shall present those principles which are of most concern to businessmen. Then, with these principles as a guide, we shall seek to find solutions for the problems that the businessman faces.

Although the businessman cannot afford to ignore the effect of his activities on social welfare, nevertheless this is not his chief interest. His principal purpose is, on the contrary, to maximize his own wealth and income. A knowledge of economic principles greatly increases his chances of success because it enables him to make a better adjustment of his own business to the general economic situation.

Study Questions

1. Since all factors of production are subject to risk, why is the risk taken by the businessman of such special significance?
2. It is often maintained that the consumer stands at the center of all economic activity. In what sense is this true?
3. Why must the businessman be concerned with the social implications of his activity?
4. What four important questions may be asked to determine whether the businessman is contributing to the general welfare?
5. Our economy of capitalism is often called a profit economy. Many contemporary writers refer to capitalism as a profit-and-loss economy. Is this latter characterization correct?
6. Why must a business make profits?

7. What three basic conditions are necessary if businessmen are to function efficiently and effectively in promoting the general welfare?

8. Name some of the outstanding economic problems that confront our nation at the present time.

References for Further Reading

BLODGETT, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 1.

BOULDING, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Part I, Chap. 1.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 1.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 1.

MARSHALL, ALFRED: "Principles of Economics," 8th ed., Macmillan & Co., Ltd., London, 1925, Book I, Chaps. 1-4.

2. The Foundations of Economic Activity

Our study to this point has indicated that if businessmen are to be successful in operating their enterprises they must take many factors into account. Among these, none are more significant than the three foundations upon which all the principles that govern economic activity rest. These basic foundations are (1) the nature of man; (2) the nature of his physical environment, for physical laws vitally affect his efforts to satisfy his wants; and (3) the nature and purpose of organized society. We shall briefly consider each of these foundations in order to indicate its importance to an understanding of the fundamentals of economics.

The Nature of Man

Every businessman knows that he cannot operate his business successfully unless he takes full account of the nature and characteristics of the one with whom he does business. It is essential, therefore, that businessmen recognize the fact that individuals devote a very large share of their efforts to activities that will satisfy their wants or desires. If they are hungry, they seek ways and means to secure food and then they consume it. If they become tired, they find a place to rest and then they fall asleep. Sometimes they desire action, and so they formulate rules to govern sports and then they participate in them. All these activities and many others illustrate a plain truth; namely, that individuals have many desires that they seek to satisfy. This truth lies behind most human activities, and the businessman must take it into account. In our study of economics we call this fundamental truth the *principle of want gratification.*

This basic principle, however, must be modified to some extent. Careful observation will reveal that, although individuals are seeking constantly to satisfy their wants or desires, they are seldom wholly successful in satisfying even one of them. At times it is possible to satisfy a want for the moment. However, once satisfied it returns time and again, and the process of gratifying it must be repeated. Still more significant is the fact that, although any one want may be temporarily satisfied, our wants are so numerous and varied that it is never possible to satisfy all of them even for a moment.

Our observation will reveal, further, that not all wants or desires which are to be satisfied are of direct economic concern; hence, businessmen must distinguish between those activities which are engaged in to secure goods or services that command a price in the market and those which are not. A man's effort to secure and retain the love of his family or esteem among his friends, or to satisfy a spiritual longing, may be very significant to him and yet may have no effect in the market. No matter how important a role such activities play in the lives of men and women, they fall outside the field of economic relationships. The businessman, therefore, must be sure that he distinguishes activities which do not affect the market from those which definitely are undertaken for the production, exchange, or distribution of goods and services.

The pattern of economic activity that businessmen must take into account is influenced by two other considerations. Both of these grow out of the fact that men seek to gratify their wants. The first consideration is the *principle of diminishing utility,* which may be stated as follows: *As a person acquires or consumes more of any commodity in a limited period of time, beyond a certain point his desire for additional units tends to decrease.* This principle is very important and will be discussed later.

The second factor that vitally affects the pattern of want satisfaction is that man is endowed with *foresight* and can, therefore, envision the future and plan for it in the present. Because of this characteristic, individuals often receive as much satisfaction in anticipating the consumption of a good as in the actual consumption at a later time. In this respect man is different from animals, for if they provide for the future, it is instinct and not foresight that guides them. Foresight causes individuals to refrain from satisfying desires in the present in the expectation of receiving a greater amount of satisfaction in the future. An individual, for example, may save part of his income in the present because he wishes

to enjoy extra income in the future. Individuals also save to protect themselves against insecurity in the future; and often they make great sacrifices in the present because they expect to receive even greater future benefits. This characteristic of foresight is very significant. On the one hand, it results in savings from which comes the capital required for the development of business enterprises; on the other hand, it sometimes forces the businessman, in seeking markets for his products, to urge individuals to spend on present wants rather than save for the future.

The fact that individuals are endowed with foresight leads to still another factor which the businessman must take into account. Because of the law of diminishing utility and because individuals can never gratify all their desires, they tend to choose between goods and services in such a manner that their satisfactions will be balanced. Once consumption of a good reaches a certain point, each additional unit gives a smaller degree of satisfaction. Therefore, the individual soon turns from the consumption of one good, the desire for which has been partially satisfied, to the consumption of another, the desire for which is still relatively intense. When the individual has partially satisfied his present desires, he weighs their further satisfaction against his possible wants in the future. By thus balancing one desire against another, he tends to spend his income in such a way as to give him the greatest possible total satisfaction. If an individual's wants were in perfect balance, a dollar spent on one good would yield exactly the same satisfaction as a dollar spent on any other.

In view of these facts, the student of economics should keep in mind three basic principles relating to man: (1) Man is endowed with a great many desires that he seeks to satisfy. (2) In gratifying these desires he discovers that the principle of diminishing utility operates. (3) He receives the greatest total amount of satisfaction if he balances his desires, both in the present and in the light of the future.

The Nature of Man's Physical Environment

The second foundation for the study of business economics is a knowledge of man's physical environment and the limitations it places on his ability to satisfy his wants. In some parts of the world nature has been bountiful, in others niggardly. Soil, climate, and natural resources

are all vital factors in determining the success or failure of man's economic efforts.

Whether a country is rich or poor in natural resources, there will always be some goods and services that are relatively plentiful and others that are relatively scarce. Moreover, as population increases, farm lands and other natural resources (aside from new discoveries) tend to remain constant or even decrease, so that less on an average is available for each individual. It follows, therefore, that at any given state of technological development, the additional goods required for the additional people can be produced only at progressively greater expenditure of labor and capital. This situation will result because a principle known as the *law of diminishing returns* is operating. This law states that *as additional units of labor or capital are applied to a fixed amount of land, beyond a certain point there will not be a proportional increase in the product.*

In a later chapter this principle will be explained more fully. Here it is introduced merely to illustrate the limitations that man's physical environment places on his ability to satisfy his wants.

The Nature of Organized Society

The third foundation for the study of business economics is the nature and functions of organized society. As we know them, economic or business activities are always social. To carry them on, individuals must organize, and they must have various kinds of dealings with other people. If he were not a member of a social group, man could satisfy very few of his wants. A knowledge of the nature and pattern of society is, therefore, indispensable to an understanding of man's economic activities.

Relation of the individual to society. When human beings live in social groups, how do the economic interests of the businessman relate to those of the group as a whole? On this question two extreme points of view are possible. The first is that the interests of the businessman inevitably conflict with the social interest. Therefore, since the social interest is of paramount importance, the businessman should subordinate his economic activity to the needs of the group. If he refuses to do so voluntarily, society should restrict and control his activities in the interest of the common welfare.

The second extreme point of view is sometimes called the *doctrine of economic harmonies.* It holds that the interests of the businessman are

in complete harmony with those of the group. Whatever a businessman does to gain wealth and income for himself also benefits society, and this is true even though he does not so intend it. It follows, therefore, that any regulation of business activity is undesirable. According to this theory the businessman is the best judge of his own interests; and when he is allowed to follow his own interests without restraint, he also makes the greatest contribution to the social welfare.

It should be clear to the student that, though both of these theories represent extreme points of view, they both contain some element of truth. Actually, man's business efforts represent neither an inevitable conflict with the public welfare nor an unfailing means of serving it. Many examples could be cited to show how individuals through monopoly, fraud, or other means have been able to enrich themselves at the expense of society. However, even more examples could be cited of profitable business activities which result in social benefits.

In spite of the fact that business activity based on self-interest cannot categorically be classified as either harmful or beneficial to society, governmental policies have often been based on the acceptance of one or the other of these extreme views. In the United States the doctrine of economic harmonies was for many years accepted with few reservations. Even today it remains a basic tenet of our philosophy of government, although in practice it has been greatly modified. The wide adoption of this doctrine was due partly to the fact that conditions in the late eighteenth and the nineteenth centuries were conducive to faith in the social benefits of business activity; and partly to the clear statement of the doctrine by one of the greatest writers in the history of economic thinking. This writer was Adam Smith, the Scottish philosopher, who is often called the father of economics.

Adam Smith wrote a treatise on economics called "The Wealth of Nations," published in 1776. This treatise was, among other things, a protest against the many rules and regulations that states had imposed on business. Smith argued that business activity is socially useful. In a much-quoted passage in his book he maintained that the businessman, though he may be seeking only his private gain, is "led by an invisible hand to promote an end which was no part of his intention." The end to which Smith referred was an increase in the national wealth. By the "invisible hand" he meant the force of competition in a free market.

Smith was too able a man to accept completely the doctrine of

economic harmonies. Nevertheless, his emphasis upon the fact that the interests of businessmen and of society are parallel caused widespread demands that the government cease its policy of interfering with business. Later the demand that government regulation of business be held to a minimum came to be known as the doctrine of *laissez faire.*

Since Adam Smith's day vast economic changes have taken place, and some of these changes have increased the need for government control of business. In some industries competition has almost disappeared, while in others it is much less effective than formerly. Contributing to the decline of competition are mass-production manufacturing, the formation of great combinations, and the development of the modern public utilities.

Today few people question the need for a considerable amount of government control of business. It is generally recognized, for example, that the public utilities should be regulated, since they must operate as monopolies to be even reasonably efficient in serving the public. However, aside from the utilities there is considerable disagreement among economists as to *how much* and *what kind* of controls are needed. Any regulations imposed should, of course, be for the benefit of society. In general, government regulations that open the channels of trade and increase economic opportunities benefit nearly everyone. On the other hand, both business and the public welfare are likely to be injured by regulations that hamper initiative or put unnecessary restrictions on production and consumption. The view most generally accepted by those who believe in the basic doctrine of economic harmonies is that, whenever business is subjected to controls, the government should formulate a definite policy and should make clear its objectives.

An enlightened program that may be followed by government in regulating economic activity has been set forth by Professor J. M. Clark. He states that when governmental controls are imposed on business the following procedures should be employed: First, define the objectives with as much precision as possible so that all persons who will be affected by them may receive the greatest possible degree of consideration. Second, decide on the kind of pressure that will be used in attaining the objectives. For instance, will the government insist that the business affected operate only according to certain well-defined rules, or will it simply prohibit a certain business practice? Third, make the controls in accordance with the prevailing political ideals. For example, we speak

of our government as a democracy. Democracy, however, implies freedom; and if our economic system is to remain free, only the most necessary controls should be imposed on business.

Professor Clark has given eleven rules or tests for a reasonable system of control, and they are essential to an understanding of this problem. He writes:

> A good system of control must meet a number of tests, some of them quite difficult. (1) It must be democratic. . . . (2) It should know what it wants. . . . (3) It must be powerful—powerful enough to make an unwilling minority obey the will of the majority. . . . (4) It must be efficient and at the same time must not destroy the efficiency of the thing it is regulating. . . . (5) It must "economize coercion." (6) It must utilize all of the strongest and most persistent motives of human nature, both generous and selfish: hope of reward, fear of punishment. . . . (7) The duties imposed must be simple enough to be understood, and this means, among other things, that social control must follow precedent a good deal of the time. (8) Control must be guided by experience or be wisely experimental. (9) It must be adaptable. (10) It must be farseeing. . . . (11) And lastly, social control must be capable of progressively raising the level of mankind.[1]

To the extent that the government in regulating business adheres to these general principles, the system of free enterprise may be used to enhance the general welfare. In the succeeding chapters we shall assume that the government will follow these rules laid down by Professor Clark. We shall further assume that, wherever society has found interference with this freedom to be necessary, the principal purpose of such interference has been to restore equality of opportunity in the market, so that both producers and consumers can be assured the greatest amount of satisfaction.

Three Characteristics of Modern Economic Society

We have pointed out that in a modern economic society like ours individuals are free within broad limits to engage in whatever business they choose and to carry it on as they wish. Such a system is called *capitalism* or *free enterprise*. In a system of free enterprise we find three

[1] Clark, J. M., "Social Control of Business," McGraw-Hill Book Company, Inc., New York, 1939, p. 13.

basic characteristics, which to some degree conflict with one another: *voluntary cooperation, competition,* and *monopoly.*

Voluntary cooperation. One of the outstanding characteristics of free enterprise is the degree to which individuals cooperate to produce what they want. Individuals anticipate that certain types of goods will be in demand and so they make plans to produce them. Often they must obtain the aid of others. If a product is to be manufactured, it may be necessary to form a corporation. Capital must be secured, buildings constructed, raw materials purchased, labor hired, and markets found. The whole process of production and distribution requires cooperative arrangements with other individuals. In this process of cooperation there must be division of labor, and in general each individual will seek to perfect himself in that work which he can do best and which will bring him the greatest return.

Competition. In spite of the fact that much voluntary cooperation is found throughout our economic system, there is also a wide field in which competition prevails. To make profits a businessman must market the product he produces, but there are always others who are trying to market either the same product or substitutes for it. No matter how great is his control of the market for a product, competition in some form and to some degree will always be present. In a free economy no one is required to purchase from anyone else; and if the price of a good is not reasonably satisfactory, consumers are sure to encourage the production of substitutes. Competition, consequently, is a power that forces businessmen to produce what consumers desire at a price that they are willing to pay.

Monopoly. Though competition in one form or another pervades all business activity, it has already been pointed out that monopoly is becoming increasingly important in some parts of our economy. Usually when we speak of monopoly we mean the situation that exists when there is only one seller of a product in a given market. Less frequently we mean that there is only one buyer. The power of a monopolist depends largely on whether there are good substitutes for his product. If the substitutes are good, a monopolist's power over price is small; if they are poor, his power may be great. Monopoly or near monopoly is found in various industries. In one very important group, the public utilities, it is a dominant characteristic. In any case, monopoly is so significant that it must be regarded as a basic characteristic of our economic system.

Competition and Monopoly Compared

Having discussed the basic characteristics in a system of free enterprise and noted that both competition and monopoly are significant, we must now compare the advantages and disadvantages of these two factors.

Advantages of competition. Competition is such a significant force in our economy that its advantages and disadvantages may well be considered in detail. In fact, the decline of competition in the American economy and the development of powerful monopolies have been regarded by some economists as an ominous threat to our free-enterprise system. Recently the following advantages of competition were advanced in a study made for the Temporary National Economic Committee (TNEC), Seventy-sixth Congress: [2]

1. Since the resources of the nation, land, labor, and capital are scarce and since a variety of goods may be produced, "economy requires that scarce resources . . . be allocated among the nation's industries" to produce those goods which consumers demand most. Competition, if allowed to operate, will produce this result. If there is freedom of entry into business and if the resources of the nation are mobile, land, labor, and capital will be drawn from fields in which the demand is not great and added to other fields in response to increased demands.

2. Competition serves the consumer by protecting him against extortion. As long as the consumer has alternatives, he can prevent the producer from taking unfair advantage of him.

3. Competition enhances the quality of the good and reduces prices, for if the producer wishes to increase his sales in order to increase profits, he must offer the consumer more goods for less money.

4. Competition leads "to the continuous improvement of industrial efficiency." When producers are in competition, any improvement adopted by one will give him an advantage. Consequently, each producer is constantly seeking to become more efficient, for as soon as one producer adopts an improvement, all others must quickly follow suit or they will lose out.

[2] Wilcox, Clair, "Competition and Monopoly in American Industry," Monograph 21, Temporary National Economic Committee, Senate Committee Print, Washington, 1940, p. 15.

5. Competition makes for material progress. As pointed out above, each producer must constantly seek new and better ways of operating his business. This fact means more goods at lower prices.

6. Competition often operates slowly and may inflict hardships upon producers; nevertheless it ultimately serves the common good, for it causes businessmen to increase output and provide more goods at lower prices. The results are fuller employment, improvement in technology, higher wages, lower prices, more goods, and a wider distribution of goods. Consequently the standard of living is raised to the highest possible level that productive resources and technical skill can maintain.

Disadvantages of competition. The members of the committee also discussed the disadvantages of competition, since they discovered that competition leads to certain types of waste. They list the following disadvantages: [3]

1. Competition calls forth a needless variety of models and sizes, and places undue emphasis on style and fashion. This condition causes competition "in persuasion" rather than "competition in service." In other words, quality may be sacrificed for style.

2. Although competition may result in higher wages, shorter hours, and improved working conditions, this result does not always follow. Competition may have the opposite effect; for example, managers in order to compete may seek to lower costs by cutting wages and lengthening hours. The one great difficulty in a competitive system is to keep the plane of competition sufficiently high.

3. Competition may cause "inefficiency in the utilization of natural resources." This result may occur because, in order to lower costs, the short-run rather than the long-run point of view may be taken. Consequently, wasteful methods of cutting timber, mining, and caring for the soil may be employed. Thus, "competition is not conducive to conservation."

4. Competition often leads to an unwarranted "duplication of plant, equipment, and personnel." The surplus of oil stations, grocery stores, and similar small businesses is ample evidence of such unwarranted duplication.

5. Competition "makes for secrecy and impedes the communication of new ideas," for each producer wishes to retain any advantages that he may have discovered.

[3] *Ibid.*, p. 15.

6. Competition often prevents "the coordination of services." In certain fields, goods or services can be more efficiently rendered if the complete process is under one control.

7. Competition is not without its costs. It may require a high rate of business mortality; it may inflict serious losses on investors. It may deprive workers of employment, and it may create many other losses.

Thus, it can be seen that the advantages of competition in some fields may be partially offset by disadvantages. This fact indicates that competition may well be allowed in some fields but curtailed in others. In fact, if our economy is to operate efficiently, we cannot hope to rely entirely on competition. When free competition produces some of the evils mentioned above, the general welfare is often promoted by allowing producers to operate as monopolists.

Advantages of monopoly. The principal advantages of monopoly are the following:

1. Wasteful duplication can be avoided.
2. Sales costs can be reduced.
3. Technological progress can be faster, since new discoveries and inventions may be used immediately by the entire industry.
4. There is less pressure to reduce wages or to adulterate the product.
5. The investor can be given greater security, and bankruptcy can more often be avoided.

All the above advantages are in the public interest, and in industries where competition does not produce still greater benefits monopoly should be allowed.

Disadvantages of monopoly. Monopoly possesses disadvantages, however, which frequently more than offset the advantages. The report of the TNEC points out that the following ten indictments have been leveled against monopoly: [4]

1. Consumers may not be allowed to secure what they desire but must take what the producer makes.
2. Consumers may have no protection against high prices, since the monopolist controls supply and he can charge any price he desires simply by bringing less goods to the market. Moreover, the monopolist does not need to worry about quality.
3. The existence of monopoly is no guarantee that the worker will be protected. Since the monopolist may be the only important employer of

[4] *Ibid.*, pp. 17-18.

labor in a certain locality, he may be able, if he so desires, to establish low wages, long hours, and undesirable conditions of employment.

4. The community has no way to punish the monopolist if he is inefficient. Under competition he must become efficient or lose out.

5. Although the monopolist may discover new techniques or invent new products, society has no way of forcing him to use his discoveries as soon as they are available. Progress requires that old techniques and old machines be scrapped. Since the monopolist is interested primarily in profits, he may find greater profit in holding back new discoveries than in using them.

6. Monopoly may not lead to the full utilization of the plant. Here again, desire for high profits may cause the monopolist to withhold goods from the market.

7. The monopolist may not adjust his prices to changes in the market. Under competition, prices change as demand changes. In periods when demand declines, competitors are forced to lower prices, but the monopolist does not need to do so. This failure to adjust production and prices to changes in demand may lead to unemployment, which in turn will make the whole economy unstable.

8. Many of the foregoing practices used by the monopolist will prevent the raising of the standard of living. If prices are high, the product is of poor quality, wages are low, hours of employment are long, and output is limited, then the standard of living cannot be raised.

9. Since the monopolist can determine prices, wages, and also profits, he can determine the share that is to go to workers, investors, and consumers. He can thus make this distribution very unequal, for he may exploit any one or more of these groups for the advantage of the others.

10. The monopolist by abusing his power may cause the state to impose very restrictive rules. Consequently private enterprise may be so restricted that to all intents and purposes it gives way to public enterprise. Under such conditions the system of private enterprise would be destroyed and our democratic system would be abandoned.

In the light of the foregoing facts, monopoly clearly has its disadvantages, as well as its advantages, just as does competition. However, in some industries monopoly is desirable. The problem is to make sure that competition is protected where it offers great advantages, and that, on the other hand, monopoly is permitted where competition is unsatisfactory. Wherever monopoly is allowed, however, the state must see that

its advantages are secured and its disadvantages minimized. In our present economy both competition and monopoly exist; monopoly, however, is generally placed under state supervision.

Between the extremes of free competition and monopoly there are many businesses that require restriction to some degree in the interests of the general welfare. Such restrictions may simply encourage competition by keeping the channels of trade open; in other cases they may encourage production of certain types of goods that would otherwise not be produced at satisfactory prices. Again, society may consider that certain goods, such as opium, are detrimental to its welfare, and it may prohibit their production or sale. Then, too, the state may insist that goods be produced under certain conditions of sanitation, and also that they be labeled correctly. Hours and conditions of labor may be carefully supervised. Moreover, if anyone were allowed to enter into any occupation he desired, some who were unfit might carry on activities that were not in the interest of the general welfare. In this case the state, to protect the general welfare, is likely to restrict freedom by indicating the conditions under which individuals may enter certain occupations. Thus doctors, lawyers, teachers, and some others must be certified before they are allowed to practice their professions.

Many other limitations of competition are found in a system of free enterprise, but even so competition is still the most vital force that operates in the market.

Summary

In this chapter we have been concerned with the three basic foundations of economic activity, which businessmen must take into account if they are to operate their enterprises successfully: (1) the nature of man, (2) the nature of his physical environment, (3) the nature and purpose of organized society.

The discussion revealed that businessmen are often restricted in their activities and that unless they are aware of and take into account these restrictions, they may fail in their objectives. Thus, the businessman must understand the *principle of want gratification.* However, he must realize not only that many of man's wants are of no significance so far as they affect the market, but also that those wants which are significant follow the principle of diminishing utility. Further, because man has foresight and seeks in consequence to balance his desires, businessmen must do two things: On the one hand, they must at times encourage saving in order to secure the necessary capital to

expand their plants; and on the other hand, they must at times overcome sales resistance by convincing the consumer that present consumption will bring greater satisfaction than future consumption.

The businessman must also understand that there are many limitations to increasing production. Of these none is more significant than the niggardliness of nature in supplying natural resources. Businessmen, therefore, must constantly seek to overcome the *law of diminishing returns.*

Finally, the businessman must take into account the objectives and purposes for which individuals organize. He must understand the extent to which there is inevitable conflict and the extent to which there is harmony between the economic purposes of individuals and those of society. In the United States the *doctrine of economic harmonies* has been accepted, though modified to suit changing needs; and much faith has been placed in the power of free competition to control business in the public interest. Our government, however, has not hesitated to impose controls on enterprises that are best operated as monopolies.

Three characteristics play an important role in modern economic society: voluntary cooperation, competition, and monopoly. The function of government is to encourage voluntary cooperation while at the same time determining whether such cooperation is to operate under the rules of free competition or under conditions of controlled monopoly. Since competition and monopoly are such potent forces, the advantages and disadvantages of each were considered.

It is now necessary to understand certain basic terms before we return to the discussion of economic principles. These terms will be defined in the following chapter.

Study Questions

1. Upon what basic foundations does all economic activity rest?
2. State the principle of want gratification.
3. What other considerations relating to want gratification are of significance to the businessman?
4. Distinguish between the principles of diminishing utility and diminishing productivity.
5. State two basic doctrines that concern man's economic activities as they relate to society. Which point of view is accepted in the United States?
6. What contribution was made by Adam Smith to economic thinking?
7. What is the doctrine of *laissez faire,* and what is the role of government when this doctrine prevails?

8. Does the laissez-faire doctrine prohibit all regulation by the state of the activities of businessmen?

9. State Professor Clark's rules on tests for a reasonable program of regulation.

10. Study some law passed by Congress in the light of these rules, to determine the effectiveness of this law in promoting the general welfare. Some laws that may be considered are Wages and Hours Act; Securities Exchange Act; Sherman Antitrust Act; Securities Act; National Labor Relations Act; Public Utility Holding Company Act.

11. Name the outstanding characteristics of modern economic society.

12. What are the main advantages and disadvantages of allowing free competition?

13. Since in our system both free competition and monopoly prevail, what rule must be laid down to determine the conditions under which each is to be allowed to prevail?

Exercises

1. Several years ago an editor of a farm journal stated:

> We have, and our folks should have, the utmost confidence in economic law, which is only another name for human nature and behavior. . . . We retain complete confidence in economic law. It is irresistible, and the laws of Congress must yield to it, as they have many times in the past. Congress does as it pleases, but nothing it does in defiance of human nature and behavior will work or can be made to work. It is a pity so few Congressmen understand that.

a. Is the sentiment expressed in this statement similar to that stated by those who believe in *laissez faire?*

b. In what respects do you agree with this statement? Disagree?

2. The following advertisement appeared in a leading magazine. Under the heading "The goose that lays the golden eggs is beginning to realize she is a goose," the following statements are made:

> American industry—spurred on by the fight for profit—has paid more taxes, provided more jobs, paid higher wages, produced more goods at lower prices, developed more natural security than any other system in any other time in any other part of the world.
>
> Now this great force, which has done all this for America, is being treated by elements of government and labor as though it were an enemy. Industry is

told what it must pay but government wants to decide what it can charge:—a thousand regulations trip, hamstring, and block business progress. The wonder is that American industry and those who own it keep on hoping and working.

a. In what system of government does the author of this statement believe?

b. Do you agree that business is being "hamstrung"? Why or why not?

c. Would Professor Clark's rules, if applied, cause a similar statement from business? Why or why not?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 6.

Clark, J. M.: "Social Control of Business," McGraw-Hill Book Company, Inc., New York, 1939.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 3.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 3.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 3.

Hamilton, W. H.: "Pattern of Competition," Columbia University Press, New York, 1940.

Wilcox, C.: "Competition and Monopoly in American Industry," Monograph 21, Temporary National Economic Committee, Senate Committee Print, Washington, 1940.

3. Fundamental Concepts

This chapter will be devoted to the definition of certain terms that are used frequently in the study of economics. One difficulty that confronts the economist in his explanation of principles is that most of the terms he employs are in everyday use. As a consequence most individuals assume that they understand what the economist is seeking to explain. Often, however, they do not, for a given term may have a wide variety of meanings. In order, therefore, to be sure that the reader understands what a term means, the economist must define it in a special or technical sense.

Good. A *good* is anything that is desired by someone. It may also be defined as anything that has the power to satisfy a desire. Goods may be material, or they may consist of want-satisfying services.

Utility. The power that any good possesses to gratify a desire is called *utility*. All goods, therefore, possess utility.

Free goods. Some goods that are desired are available in such abundance that all may use them without cost. Such goods are called *free goods*. Air is an example. In some instances water is also free. Free goods possess utility, but are not scarce and are not owned by anyone.

Economic goods. These are goods which are available in such limited quantities that those who wish to enjoy them must pay something for them. Economic goods, therefore, possess utility and scarcity; and they can be transferred from one person to another.

Services. Economics is concerned with many nonmaterial goods that satisfy desires. These nonmaterial goods are called *services*. The doctor, lawyer, teacher, musician, servant, and actor provide services. In many cases nonmaterial economic goods or services may be more important than material goods.

Wealth. If an inventory of all material economic goods in the United States could be taken at one time, we should have an idea of the *wealth* of the nation. Wealth consists of a stock of material economic goods

which is in existence at any one time. It should be noted that all wealth is owned. Further, it should be noted that wealth consists of the goods themselves, and is not a monetary evaluation of the goods. Because all of us are so used to valuing goods in terms of money, we are inclined to think of money as synonymous with wealth. The objection to thinking of wealth as a monetary evaluation of goods becomes clear during periods of rapid changes in the price level. A house, for example, which may be sold for $10,000 when prices are at one level, may be sold for $20,000 a few years later when the price level has increased; or if the price level has decreased, it may be sold for only $5,000. Yet it is the same house and can gratify the same desire whether its price is $5,000 or $20,000. The wealth of a country, therefore, is the stock of goods that it possesses rather than the monetary value of these goods.

We have included in the term "wealth" only the stock of material goods. Some might argue that we should also include the abilities and skills of people. Does not the skill of a doctor, for example, have great economic value? The trouble, however, with including skills in the category of wealth is that they cannot literally be exchanged and therefore have no definite and determinable economic value. The only way in which a doctor can sell his skill is by selling his services. These services are not wealth because they are not material. They do, however, have a definite market value, and we have recognized this fact by classifying them as nonmaterial economic goods.

Producers' and consumers' goods. Some economic goods satisfy consumer wants directly; others are wanted, not for their immediate utility but to aid in producing other goods. A good that is directly consumed is called a *consumers' good.* Any good that is used to advance the process of production is called a *producers' good,* because it has utility only indirectly. The distinction between a consumers' and a producers' good is often indefinite. For example, coal used to produce steam for a factory power plant is a producers' good, but coal used to heat the home is a consumers' good. It is plain, therefore, that the purpose for which a good is to be used determines whether it is a producers' or a consumers' good.

Production. Any activity that contributes directly or indirectly to the satisfaction of economic wants is productive. Hence, *production* is the creation of utility. Production may consist of putting a commodity into a more desirable form, or creating *form utility.* It may provide a good

at a more desirable time, creating *time utility*. It may bring it to the place where the consumer can use it, creating *place utility*. It may put it in the possession of the consumer, creating *possession utility*. Finally, it may consist, not in dealing with material goods at all, but in giving direct personal services. In this case it is said to create *service utility*.

Factors of production. Anything that contributes to the productive process is called a *factor of production*. Factors of production are usually classified into four large groups: *Land* is all natural resources. It includes not only farm land but also forests, mines, water-power sites, and all the other gifts of nature. *Labor* is human services. *Capital* is all produced material goods that are intended to aid in further production. It does not, of course, include consumers' goods. *Entrepreneurship* is that special type of human service which is concerned with organizing and directing business activities. Entrepreneurship involves not only controlling business but also taking the major risks.

Money. As has been pointed out, wealth or the stock of economic goods is usually measured in terms of money. *Money is merely a commodity or token that is used as the common medium of exchange and the common measure of value.* It has been pointed out that money is a good measure of wealth only in periods in which the price level is reasonably stable. In a later chapter we shall discuss the characteristics of money in some detail. The term is introduced at this point only to point out that money is not wealth but simply a device used to exchange and measure wealth.

Income. Wealth is a stock of economic goods in existence at a given moment of time. Income, on the other hand, is the flow of economic goods to consumers over a period of time like a month or a year. Wealth and services are being constantly produced and consumed. An individual, for example, works to produce goods for which he receives a wage or salary. He then spends the money that he has received to purchase the goods that he desires. The goods so purchased constitute his real income. Income, then, should be thought of as consisting of *a flow of economic goods over a period of time, such as a month or a year.*

Types of income. Though the basic meaning of income is a flow of goods, income may be considered from three points of view. First, there is *money income*, which is the amount of money one receives during a given period of time like a week or a year. Second, there is *real income*, which consists of the goods or services that money income can buy. Real

income will increase if money income remains constant and the prices of goods and services decrease. On the other hand, real income will decrease if money income remains constant and prices increase. Third, there is *subjective income,* or the amount of satisfaction that a person receives in consuming goods and services. Since different people have different wants and different powers of enjoyment, two individuals in the same occupation who are receiving the same salary may receive different amounts of subjective income. However, since subjective income cannot be measured, it is a less useful concept for economic analysis than either money income or real income.

Property. Property is defined as the exclusive right to possess, enjoy, or dispose of an economic good. Thus property consists of a bundle of rights which an individual can exercise by the fact that he owns a good. These rights, however, can be enjoyed only to the extent that the owner is protected by the state. Since property consists not of goods but of rights in goods, the state may modify these rights at any time. Suppose, for example, that an individual owns a piece of land in a city. Since he owns the land he has the right to use, enjoy, or dispose of it as he sees fit, subject only to restrictions by the city. Suppose the city decides to pass a zoning ordinance in which it limits the use of this land. It is obvious that the individual's rights have been modified; yet he still owns the land. It should be clear that, although all economic goods are owned, the rights of use and enjoyment may vary from time to time and from place to place.

Public versus private property. It was pointed out that all wealth is owned. However, the community as a whole represented by a governmental unit may own a particular piece of wealth. Thus, cities own parks, roads, water systems, schools, electric or gas plants, street railways, sewage disposal plants, hospitals, and many other kinds of property. In the United States vast amounts of land are publicly owned, as well as lakes, rivers, canals, and hydroelectric developments. Thus it may be seen that ownership of property by individuals may become relatively less significant if governmental units extend ownership to include more goods. In spite of the increase in public ownership, private ownership is still the rule in this country.

Value, utility, and price. People often confuse the terms value, utility, and price. In economic usage value is strictly an exchange concept in that the value of a good is the other goods that can be obtained for it in

trade. Thus, the value of a suit is its exchange power in terms of wheat, gallons of gasoline, or pairs of shoes. The utility of the suit, however, will vary with the needs of the individual at the moment.

The value of a good, then, is its power to command other goods in exchange. Price, on the other hand, is the value or exchange power of a good stated in terms of money. In other words, the price of a good is simply the amount of money that can be obtained for it in exchange.

Summary

A GOOD is any object or service that has the power to satisfy a want.

UTILITY is the want-satisfying power that any good possesses.

FREE GOODS are goods which exist in such abundance that all may use them without cost.

ECONOMIC GOODS are goods that command a price. Besides possessing utility, they are scarce and transferable.

SERVICES are all nonmaterial goods that satisfy desires.

WEALTH is a stock of material economic goods in existence at a given moment of time. Wealth is always owned.

CONSUMERS' GOODS are goods that satisfy consumer wants directly.

PRODUCERS' GOODS are goods that do not satisfy wants directly but are used to produce other goods.

PRODUCTION is the creation of utility.

FACTORS OF PRODUCTION. The four factors of production are *land*, *labor*, *capital*, and *entrepreneurship*. Land is all natural resources; labor is human services; capital is produced material goods intended for use in further production; entrepreneurship is organizing and directing business activity and taking business risks.

Study Questions

1. Why should the fundamental concepts that are used in economics be carefully defined?
2. What is the difference between a good and a service?
3. What is the difference between a free good and an economic good?
4. Under what conditions may a producers' good become a consumers' good? Name a good that might be either.
5. Define utility. Name the various kinds of utility explained in the text.

6. Compare the definition of production given in the text with that commonly used when the term is used. According to the definition in the text, is transportation production? Why?

7. Name the chief factors of production.

8. Name the types of income mentioned in the text and define each.

9. Distinguish between economic goods and property.

10. Distinguish between value and price.

Exercises

1. After the First World War a gum manufacturer used the following advertisement: "Five cents before the war, five cents during the war, and five cents after the war." In view of the great changes in the price level, did the value of this commodity change during the periods indicated?

2. In England the government has taken possession of several basic industries. Did this change in ownership change the wealth of the country? Could it change the amount of producers' goods? Of consumers' goods? Did the amount of property in existence change? Explain carefully.

3. Between 1928 and 1932 prices of many commodities decreased. If the amount of goods in existence in 1932 and sold at low prices was exactly equal to the quantity in existence in 1929, did the wealth of the country increase, decrease, or remain the same? Why?

4. An advertisement stated, "Value is the combined properties of a thing that make it useful or desirable. Price is the quantity of one thing that is demanded or exchanged in barter or sale for another." Rewrite this statement to make it mean what the advertiser intended it to mean, in terms of the definitions of price, value, utility, and so forth, as stated in the text.

5. Suppose that all goods were free. Would their utility and their value be the same as now? Explain.

References for Further Reading

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chaps. 1-3.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 1.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 1-3.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn and Company, Boston, 1937, Chaps. 2 and 3.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chaps. 1 and 2.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chap. 1.

MARSHALL, ALFRED: "Principles of Economics," 8th ed., Macmillan & Co., Ltd., London, 1925, Book II, Chaps. 1-4; Book III, Chaps. 1-6.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. I, Chap. 4.

4. The Organization of Production

In the preceding chapters a basis was established for studying the principles that control the production, exchange, and distribution of economic goods. Economics, as we have pointed out, deals with these activities. In this chapter our purpose is to examine the nature of production and the way in which it is organized.

Our problem is this: What basic principles should the businessman employ if he wishes to maximize his profits by securing high production at low unit costs? Before answering this question, we must understand the nature of production and management; the relative economies of large-scale versus small-scale production; and the different forms of business organization.

Nature of production. In the preceding chapter production was defined as the creation of utility. Since utility is the power to satisfy a want, any activity that increases the want-satisfying power of material goods is productive. Further, any activity or service which in itself has the power to satisfy people's wants is also productive. Thus it is clear that we must regard as producers not only farmers and factory workers but also actors and singers, as well as professional people like doctors and lawyers. Since productive activities satisfy wants, people are willing to pay for them; and the simplest test of whether any type of human effort is productive in the economic sense is to ask the question, Does it command a price in the market?

As we have already seen, an activity may be productive and profitable and still not contribute to the general welfare. Such an activity should not be condoned if there is general agreement that it is undesirable. However, it is not always easy to decide what promotes the general welfare, and one of the rights that consumers prize is freedom to make

their own choices. Some products may do more harm than good, as is the case with many drugs. Nevertheless, as long as individuals want a good and as long as the state permits its use, those who provide it must be regarded as producers in the economic sense. They follow the same principles and are treated by the state in the same way as the producers of any other good.

In Chap. 3 it was noted that there are five types of utility, corresponding to five ways in which production contributes to the satisfaction of consumer wants. For a material good to have utility to a given consumer it must be in the right form, in the right place, and at the right time. This was expressed by saying that a good must have *form utility, place utility,* and *time utility.* It must also have *possession utility,* because the potential consumer cannot use it unless it comes under his control. These four kinds of utility we associate with material goods. A fifth kind is that possessed by direct personal services, which we call *service utility.* At this point we shall consider these various types of utility in more detail, in order to point out why they are of special concern to businessmen.

Form utility. A manufacturer is principally concerned with processing raw materials, that is, combining them and putting them in a form that is desired. The automobile manufacturer, for example, creates out of iron, steel, rubber, glass, and other commodities a vehicle that individuals desire. That he may not carry out all of the processing from mining or gathering the raw materials to making the parts for the final assembly makes no difference. He is manufacturing so long as he is changing the elements that form the product. Thus, we may say that a manufacturer is primarily concerned with creating *form utility.*

Place utility. The automobile manufacturer may operate his factory in Detroit and the prospective consumer may live in California. Certainly, an automobile in Detroit is of no use to the man in California. Hence, if the car can be transported to California, it will have utility there. Consequently, industries engaged in transporting a good in order to give it more usefulness are creating *place utility.*

Time utility. In addition to being in the right form and in the right place, a good must be where it is wanted at the right time. Ice in winter may be relatively abundant, but it may be scarce in summer. As a result, persons who preserve the ice until summer are creating *time utility.* Similarly, many other commodities, among them eggs and fruits, have their utility increased by being stored or preserved at seasons when they

are relatively plentiful and kept until a time of year when they are relatively scarce. In the case of still other goods utility may increase with time over very long periods, as is the case with wines, antiques, and some paintings.

Possession utility. Some industries not only aid in creating place and time utility but also create possession utility. The merchant who sees that goods are on the shelf in the right form and at the right place at the right time is also creating another utility. He aids in enhancing the utility of the goods by transferring them from the possession of producers or wholesalers to consumers. In this transfer he often takes a risk, for he must anticipate the desires of his customers. Because he has thus anticipated their desires and has enabled them to secure possession of goods when they want them, he has enhanced the utility of the goods.

Any businessman who enables the ultimate consumer to secure a good when he desires it is creating possession utility. Bankers, as well as merchants, thus create possession utility. By lending money to businessmen and consumers, they enable people to acquire goods when the goods are most wanted.

Service utility. Many individuals who are in business do not deal with material, tangible goods. Doctors, lawyers, teachers, musicians, actors, and others satisfy certain wants, but they do this only by giving direct personal services. In the same general category we may also include servants of various sorts, and people who operate gambling establishments or provide professional sports. Activities of this kind are producing service utilities. Public business activities should also be noted. Policemen, public clerks, postmen, and many others in government employment are performing want-gratifying services; as such they must be classified with those who create service utility.

The Nature and Functions of Management

The entrepreneur or businessman. Since, as we have just pointed out, production consists of many activities, someone must organize and bring the various factors of production together. Someone must also see that transportation facilities are available and must provide agencies to obtain materials and to distribute both materials and finished goods. The entire production process from the securing of raw materials to the final selling and delivery of consumers' goods calls for organization.

Those who assume the responsibility of creating and carrying on the production process at any stage are called businessmen, enterprisers, or entrepreneurs. Upon the *entrepreneur* or businessman, then, rests the responsibility of making the business organization operate smoothly. This is a very great responsibility, for the entrepreneur not only must operate the business once it has been created but often must also create the business. The businessman sees or anticipates a need, and on the basis of his observations he creates an organization. It consists of those who are willing to furnish the necessary capital to build the plant, a staff who can help in supervising the production once the plant is built, and the labor force that does the actual manual work in producing the good. The businessman must make his organization function smoothly. He must also be a shrewd buyer of raw materials and labor and a shrewd seller of his product. Upon his ability to buy and to sell at favorable times and terms will depend the success or failure of the enterprise.

The businessman, however, must be more than an efficient organizer and a shrewd buyer and seller. As already pointed out, he should be aware of the effect of his dealings upon the general welfare. The businessman who is shrewd enough to make profits in the short run may do so by taking unfair advantage of others. As a result, labor may be underpaid, owners of capital may be exploited, and consumers may be charged exorbitant prices. Sooner or later, however, these practices will be exposed. Labor, capital, and the consumers no longer will tolerate exploitation, or the state will intervene. Thus, the wise businessman will take the long-run point of view and determine upon policies that will benefit the business most over a period of years rather than at the moment.

Historically, the typical entrepreneur was a man who went into business for himself. He invested his own capital, plus any he could borrow, and he managed the enterprise. This worked very well when the typical business was rather small. Today, however, in many of our larger corporate enterprises the investment function and the management function of the entrepreneur are separated. On the one hand is a large body of investors who furnish the capital and bear the risk; on the other hand is a group of persons who specialize in operating and managing the enterprise. The first group we call capitalists, the second group managers. Both groups are performing entrepreneurial functions. However, this separation of the functions of ownership and management

often creates problems. The managers of a large corporation are sometimes more interested in operating it for their own benefit than for the benefit of the stockholders, although this is exceptional. Moreover, officers of most corporations hold at least some stock.

In the main, however, those classified as entrepreneurs perform socially valuable services. They cooperate to produce the goods that the community desires. In fact, at times businessmen overestimate the desires of the public and produce such an abundance of goods that they cannot be sold at the price the entrepreneur asks for them. Thus, in constantly attempting to anticipate the needs and desires of the public, businessmen sometimes make serious errors. In later chapters we shall take note of these errors and point out their probable causes. At this point, however, our interest is centered in the fact that production cannot be carried on unless it is organized; and it is the entrepreneur or businessman who organizes production.

Size of the Business Enterprise

The way in which business is organized depends a great deal on the size of the particular enterprise. Obviously, a large business will require a more formal and elaborate organization than a small one. In modern times there has been a trend toward larger and larger business units, and this has been a very important factor in the development of modern types of business organization. However, in spite of this trend toward large-scale business, there are still many kinds of business that can be operated most efficiently on a relatively small scale. It is important, therefore, to understand both the advantages and the disadvantages of large-scale production in order to determine the best size for a particular business and the best way to organize it.

Determining the size of an enterprise. Before the businessman decides on his type of organization, he must determine whether he can produce the good economically and efficiently on a small scale, or whether in order to compete in the market he must produce on a large scale. Many businesses begin in a small way and change their organization as their growth requires. This procedure, which seems to be a logical one, often causes difficulty because the firm may wait too long before reorganizing. The businessman should study his product carefully and, if possible, choose that type of organization which will be best in the long run.

An outstanding problem confronting the businessman is to determine the relative economy of small- versus large-scale production. The nature of the product will aid him in his decision on this point. Thus, if a product requires much labor and little capital, the chances are that it can be produced in small amounts at as low a unit cost as in larger amounts. Under this condition the businessman will probably find that most makers of the good produce small amounts, and the industry is characterized by small-scale production. If, however, great amounts of capital are needed and if with every increase in capital the manufacturing process can be refined so that unit costs are reduced, then the large-scale type of business organization is desirable.

Advantages of large-scale production. Several reasons may be given to show why large-scale production is often more economical than small-scale production. First, sometimes the buying or the selling of goods on a large scale results in economies. The volume of business may be so great that those who sell to the producer will be glad to quote him favorable prices. Large-scale selling may also reduce unit selling costs. These savings in buying and selling may create such an advantage to the firms employing them that they can undertake to sell at lower prices than smaller competitors and still make a profit. The great mail-order houses, chain stores, and similar merchandising establishments operate on this principle. Manufacturers are often glad to sell to them at very reasonable prices if they will take all of the output or at least such a substantial portion of the total output that the producer can operate at maximum efficiency. Moreover, if the producer sells to one firm, many of his selling, advertising, or other costs may be reduced. The economies are therefore shared between producer and merchandiser. As a result, costs are lowered.

A second advantage of large-scale production is that labor and machinery may be specialized. In many instances a machine can be built to do one simple operation, and a sufficient volume of production allows that machine to operate at capacity. Men may also be trained for specialized tasks, and this tends to increase efficiency and output. Also specialization often leads to simplification of each particular task. This leads to a further increase in production because men become more skillful and machines can be applied to more processes. Thus, the fullest utilization can be made of the gains that come from division of labor.

A third advantage that sometimes arises because of large-scale opera-

tion is the utilization of by-products. In the packing industry large-scale production has made possible the utilization of practically every part of the animal. Indeed, only the squeal of the hog, so the saying goes, is wasted. Because of this full utilization, the various goods can be sold at a lower price than would have to be secured if by-products could not be fully utilized.

Fourth, when goods are manufactured on a large scale, time and effort can be expended in finding new ways of producing the goods, new uses, and new by-products. Many large organizations now maintain laboratories to carry on research for these purposes. Furthermore, efficiency engineers may be employed to devise better methods of processing the good, thus reducing costs. Today most new products are discovered by those who carry on research in the laboratories of large industries. Small concerns cannot afford to spend hundreds of thousands of dollars for this purpose. Thus, the large firm is often the first to bring a new product to the market, to improve it, and to produce it at a low cost.

Fifth and finally, in businesses that must have a very large plant in order to operate at all, large-scale production results in greatly lowered unit costs. Railroads and public utilities must employ large amounts of capital. Unless this investment can be spread over many units of product, the per unit cost will be extremely high. For example, suppose the cost of building the roadbed for a railroad is $50,000 per mile, and suppose that the line is to operate between two cities 100 miles apart. Obviously, the investment in the roadbed alone is very great. Assume further that this road is used by only one train a day. Clearly, since the investment is used for short periods, the unit cost of operation will be very high. Now suppose that two, three, or even more trains per day use this track. Then the investment is used close to its full capacity and unit costs of operation can be reduced. This same principle holds in many other industries that require large amounts of capital, and it makes one of the strong arguments in favor of large-scale operation. It should be noted that in these industries an increase in the amount of product is an advantage.

Limitations to large-scale operation. Although large-scale operation has many advantages, it has a number of definite limitations. We have already pointed out that some products do not lend themselves to large-scale operation with any greater degree of efficiency than is found in small-scale operation. Herein, however, is not the main difficulty. Large-

scale operation requires an unusual degree of managerial skill if the economies mentioned above are to be fully realized. Such managerial skill is not always available. Moreover, when managerial functions must be delegated, some efficiency may be lost. Then too, large-scale operation may require a vast selling organization to dispose of the product. Sometimes, also, costly advertising must be used to keep up sales. Within the plant itself, the fact that laborers may not be acquainted with management often causes them to work only as hard as they must to retain their jobs. These and other difficulties may offset much of the efficiency that could be gained if all the advantages of large-scale production could be realized. As a result, often a small producer can compete on favorable terms with a large-scale producer.

Perhaps the most basic problem confronting the management of an extremely large enterprise is this: If all decisions of any importance are made by a central authority, inefficiency may arise because of "red tape," hopeless delay, and decisions based on inadequate knowledge. On the other hand, if too many decisions are delegated to subordinates, the operations of the firm will not have the unity and coordination that efficiency demands. The management of a very large firm must try to take a middle course, but it can hardly achieve the unity, compactness, and flexibility of a well-managed firm of medium size.[1]

Forms of Business Organization

The actual number of business enterprises in the United States is difficult to determine. The National Industrial Conference Board estimated that over 10,000,000 such enterprises existed in 1935. Agriculture accounted for over 6,000,000; manufacturing, 169,000; transportation, rail, air, and motor, 6,000; trade, 1,627,000; public utilities, 60,000; construction, 70,000; finance, 120,000; and services almost 1,000,000. This last figure included 316,000 doctors and dentists, and great numbers of miscellaneous activities, such as advertising agencies, amusement places, hotels, radio stations, tourist camps, and warehouses. The Board also estimated that about 33,000,000 workers were employed in these enterprises. The data show that business is constantly undergoing changes. The year of greatest increase was 1925, when 496,000 new businesses

[1] Stigler, George, "The Theory of Competitive Price," The Macmillan Company, New York, 1946, p. 138.

began. In 1929, 483,000 were discontinued. In the years 1929 through 1937, 3,493,000 new enterprises began, and 3,591,000 ceased operating.

The data also show that owners employ three forms of organization to carry on business activity: the independent proprietorship, the partnership, and the corporation. In agriculture the independent proprietorship predominates; but in nonagricultural activities the division, so far as numbers are concerned, is as follows: individual proprietorships, 60 per cent; corporations, 30 per cent; partnerships, 10 per cent. However, these figures are somewhat misleading. In the telephone industry, for example, a total of 44,851 telephone companies is listed; yet the American Telephone and Telegraph Company and its subsidiaries account for 90 per cent or more of the entire amount of business.

In some fields, so far as volume of business is concerned, the corporation holds a dominant position. In manufacturing, corporations account for 92 per cent of the volume of the business; in transportation 89 per cent; in electric power almost 100 per cent. On the other hand, in agriculture corporations account for only 7 per cent of the volume. Thus, although the corporation does not dominate all fields, it does such a substantial amount of the total business that it may be said to be characteristic of our economy as a whole.

Because these are the three forms under which business is organized, we should study the advantages and the disadvantages of each.

The single proprietorship. Under the single proprietorship, the most common form of business organization, the owner furnishes the capital, or at least a good part of it, and assumes the responsibility of managing the business. The farmer, the butcher, the grocer, and in fact most persons conducting businesses in smaller communities both own and manage their enterprises. Hundreds of such businesses are also found in the cities. We may include under the single proprietorship most persons in the professions—doctors, lawyers, druggists, barbers, undertakers, writers, and others in similar occupations.

The single proprietorship possesses several distinct advantages. The owner can enter business or leave it as he pleases; he need consult no one. In making his decisions, he is unhampered. Since he takes the risk, he is also entitled to enjoy the profits. Such one-man control and ownership allow the greatest flexibility in organization and policy.

However, this type of organization also has its disadvantages. The owner may not be able to secure sufficient capital to enlarge the business.

Moreover, since he takes the risk he must assume full responsibility for his mistakes. Thus, he assumes unlimited liability. He risks not only what he has put in the business but also any other property that he may own. Consequently, after working for many years, a single proprietor may have his entire possessions taken from him if his business fails. Most important, as the business grows, the owner may not be able to supervise all the details; yet he must assume all the risks. Finally, no man can reasonably hope to be an expert in all phases of operating a business. The owner may be an expert in production methods, but he may not understand the problems of financing, personnel, or selling. Consequently his business may suffer. The single proprietorship is a good type of organization only if the product can be produced efficiently on a small scale. There have been cases where single proprietorships have grown into large and efficient enterprises with hired experts to handle the specialized functions of running a business, but the difficulty of accumulating capital for such enterprises has tended to restrict their growth in recent decades.

The partnership. A second form of business organization is the partnership. In such an organization two or more individuals enter into an agreement to operate a business jointly. Certain advantages accrue from such an agreement. Securing necessary capital for the enterprise may be easier, for the credit of all the partners can be pooled. Again, each partner may be an expert in a particular phase of the business; hence, the organization may secure the advantages of specialization. A partnership, like a single proprietorship, is relatively easy to form from a legal point of view. Finally, policies may be changed readily by simple agreement of the partners.

Certain weaknesses, however, are found in this form of organization. Every enterprise to operate efficiently must follow some definite policy. In a partnership it is often difficult to persuade some partners to submit to the authority of others. Without a definite policy, production may be interrupted as orders given by one partner are modified by another. The difficulty of securing absolute teamwork is a serious defect of a partnership.

A second weakness is the fact that in a partnership each member has unlimited liability so far as the firm's obligations are concerned. The actions of one partner may cause the firm to fail, and yet any one of the

partners can be forced to reimburse the creditors to the full extent of the liability. Thus, because one partner may have no assets that can be liquidated, another may have to contribute the entire deficit. This fact seriously limits the partnership, for obviously only those who are well acquainted will be willing to join in a venture that makes them so dependent on the character of their associates.

A third weakness is found in the requirement that if a partner wishes to withdraw from the firm, or dies, the remaining partners must either find someone else to take his place or purchase his share. Many times the death or withdrawal of one partner will disrupt the affairs of the firm. The partnership is not a satisfactory form of business organization in most circumstances.

The corporation. We have already noted that the corporate form of business organization now holds a predominant position. Except for agriculture, construction, the professions, and trade, every industry of importance is dominated by corporations.

A corporation is a legal entity that derives its existence from the state. To organize it, a charter must be secured; and the corporation must confine its activities to those permitted by the charter. The corporation is treated at law as being an artificial person; that is, its personality is separate from the personalities of its owners. Thus, the corporation may sue and be sued, without involving the persons of the stockholders. This separation allows, in some situations, a flexibility that partnerships lack.

When a corporation is formed, its prospective owners subscribe for stock in the company; and their voice in the management, as well as their profit or loss from operations, is based on the relative amounts of stock that they own. Generally each share of common stock carries one vote. Besides selling common stock, a corporation may secure funds by selling bonds or preferred stock, but as a rule, neither bondholders nor preferred stockholders have any voice in management. Further, their return is usually limited by contract. On the other hand, their principal is somewhat more secure and their income more assured than that of the common stockholders. They are investors and creditors rather than owners.

The corporate form of organization possesses several distinct advantages over the other two types. In the first place, each shareholder has a liability only to the extent of his investment in stock. In some corpora-

tions, stockholders are held to double liability, but even so, the amount of liability is limited. Formerly bank stockholders were generally subject to double liability, but this is no longer true except for state banks in a few states.

A second great advantage of the corporation is the relative ease of securing capital. Since there is no liability beyond the investment in the enterprise, many individuals can be found who are willing to take a risk of this kind in the hope of securing a capital gain. Moreover, since the corporation can sell not only common stock but various other types of securities, it can attract various groups of investors. Consequently it can sometimes accumulate vast sums of capital.

A third advantage of the corporation is the ease by which ownership may be acquired or relinquished. Any individual who wants to purchase stock can enter the market, and since securities of most corporations are being constantly bought and sold, he can generally secure as many shares as he desires. This flexibility is an added inducement to capitalists.

A fourth advantage is the permanence of the corporation. Individuals live and die, and if a business is dependent upon the life of one or a few persons, its plans may be disrupted. In a corporation this is not so, for even though ownership may be changing constantly, plans and policies can continue undisturbed.

Other advantages of the corporation are frequently listed. For the most part, however, they are merely incidental to the fact that the corporation can bring together great amounts of capital. Because of its almost unlimited power to obtain capital, a successful corporation can take full advantage of any gains to be had from large-scale production.

Disadvantages of corporate organization. Since the corporation is commonly engaged in large-scale production, it is often subject to all the weaknesses of great size. In very large corporations, the functions of management and ownership are often separated. Where the stock of a company is widely scattered among a great number of owners, experience shows that control falls into the hands of "management," that is, the board of directors. Often the members of the board own an extremely small proportion of the total stock outstanding. In that case the great majority of stockholders, who really own the business, are dependent on a management that may or may not operate the corporation for their benefit. In the long run a management that does not have the stockholders' interest at heart may do them great harm.

A second disadvantage of the corporation is the danger that the enterprise may be overcapitalized. The ease of securing funds and the possibility of making profits may cause those who are promoting the corporation to take an extremely rosy outlook. As a consequence, they may secure an excessive amount of capital in relation to the opportunity for immediately profitable employment of the funds. In the case of a failure of the venture when prospective profits are not realized, all who invested their funds will lose. The abuses in overcapitalization, mismanagement, and misuse of corporate power present problems that must be solved if the corporate form of organization is to serve the general welfare as it should.

Summary

This chapter has surveyed the nature and functions of business organization. The meaning of production has been explained, and the functions of the entrepreneur have been indicated. The advantages and disadvantages of small- versus large-scale production have been surveyed, as well as the kinds of business organization that best suit each kind of production. We have seen that, if the economic purpose of engaging in business is to be realized, then the proper kind of organization must be adopted and responsible management must be employed. Efficient management with a sense of responsibility will promote the general welfare; that is, it will secure for all concerned the maximum of satisfactions with the minimum of social sacrifices.

Study Questions

1. What basic problem becomes significant when economic principles are being considered by businessmen?
2. What economic services do businessmen perform?
3. What function does the entrepreneur perform?
4. Name the main economies of large-scale production.
5. What are the outstanding weaknesses of large-scale production?
6. Which form of business organization predominates in terms of number of units? Which predominates in terms of volume of business conducted? How do you account for this difference?
7. What are the main advantages of a partnership? Main disadvantages? What are the main advantages and disadvantages of the corporation?

8. Suppose you wish to enter each of the following businesses. Which form of business organization would you use? Why?

Law firm	Book publisher
Retail store	Doctor
Railroad	Manufacturer of steel
Real-estate firm	Lumber yard
Barber	Broker
Radio shop	

Exercise

The following advertisement appeared in a well-known publication:

> Who is the loneliest man in Business?
>
> The loneliest man in any enterprise is the *one* who makes the final decisions.
>
> When Allied Armies, during the war, were poised to leap on the Continent, one man, General Eisenhower, had to take his courage in his hands and pick the day irrevocably.
>
> It is the same way in business. Some one man, after all the decisions are in, has to set the course and hold to it.
>
> The consequences of his decision may affect the future of the business, the security of thousands of employees, the savings of investors.
>
> The job of making decisions—a function of management—has been well done in America. This fact is one great reason why America is the most productive nation in the world—and a country where real wages are higher than in any other country in the world.
>
> Neither the men and women in industry nor the money in industry can be effective without good management. Nothing else is so important to the workers' welfare, the investors' welfare, or the public welfare.

a. On the basis of the discussion in the text, what is the name of the *one man* who makes the final decisions?

b. Does this advertisement imply that decisions by the "loneliest man in the world" have by and large been good?

c. Should this "loneliest man" receive a reward for taking such tremendous responsibilities?

d. What kind of reward would you give him: money, honor, or respect and esteem?

e. Suppose his decisions are wrong, how would he be punished?

f. In what respects or in what way would the question of profits be related to the rewards or punishment of the "lonely man"?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chaps. 5-8.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 6.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chaps. 3-5.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 3-6.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 4.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chaps. 8 and 9.

Stigler, George: "The Theory of Competitive Price," The Macmillan Company, New York, 1946.

5. Money

No group of economic principles is more significant to the businessman than those that pertain to the determination of prices. Since prices are expressed in terms of money and since both money and bank credit are essential to modern methods of exchanging goods, it is important that we examine carefully the basic principles relating to these two factors.

The nature of money. *Money* has already been defined as *any commodity or token that serves as the common medium of exchange and the common measure of value.* The price of a good is simply the amount of money that one receives when he sells it. If no medium of exchange or money were in existence, we can rest assured that one would soon be adopted. People early discovered that the barter system would not work if they had very much business to transact. Under barter the process of trading was cumbersome, time-consuming, and therefore costly.

In the first place, it was often found that the goods to be exchanged could not be divided into units of the right value. Suppose, for example, a butcher who had chickens to sell wished to purchase an oriental rug. Obviously, he could give many chickens in exchange for the rug, and this would be satisfactory if the owner of the rug would accept the chickens. Suppose, however, the owner of the oriental rug wanted just one chicken for dinner. Since he could not cut up the rug without destroying its value, he would be forced to purchase more chickens than he desired. Some medium of exchange that can be divided into convenient sizes is therefore necessary if exchanges are to be carried on efficiently.

A second obstacle to trade by barter is that a man who wishes to purchase something may not be able to offer in exchange a commodity that the seller wants. For example, the butcher who has chickens to sell may not care to use, or even to possess, an oriental rug. As a result, the owner of the rug is unable to purchase any chickens. However, in an economy where money is the common medium of exchange, this difficulty

will not arise. Everyone is glad to take money because he can always pass it on to someone else.

Still a third obstacle to the barter system is that no common measure of value is available. In each exchange different objects are compared, shoes for cotton, wheat for corn, and the like. In any economy, however, trade becomes much easier if the exchange values of all commodities are estimated in terms of some one commodity. In a money economy it is much easier to estimate fair exchange values because everyone is in the habit of expressing the value of each commodity in terms of money. As a result, whenever there is occasion to compare the values of commodities in terms of one another, all that is necessary is to compare their money values, that is, their prices. It is clear, therefore, that money is necessary in an exchange economy.

Functions of money. It has already been indicated that money performs two primary functions: (1) to serve as a generally acceptable medium of exchange and (2) to serve as a common measure of value.

Growing out of the primary functions of money several secondary functions may be observed. Since money is the medium of exchange and the measure of value for current payments, money becomes also the measure of deferred payments; that is, I may purchase a house for $10,000 and pay $2,000 in cash, contracting to pay the remainder over a period of years. Here the exchange of a house for money has been made but the payment is partially deferred. Eventually, of course, the other $8,000 must be paid. Meanwhile, however, the price level may change, and this, as will be explained in the following chapter, is likely to create serious problems.

Another secondary function of money is its use as a store of value. Usually, money is lent or used to purchase property, and in this way income is received in the form of interest and rent. Nevertheless, at times, especially in war or in periods of business depression, considerable amounts of value or purchasing power are kept by individuals in the form of cash. This condition is not desirable, for if people who have been hoarding money begin later to spend it freely there will be a rapid rise in prices.

Monetary Standards

Gold standard. Since money is so essential in an exchange economy, it is important that the material used for money be a suitable one.

Without doubt one of the best commodities for monetary use is gold. It is durable, uniform in quality, and conveniently divisible. It has many nonmonetary uses which create a strong demand for it at all times; and it is easy to carry because a relatively small amount possesses high value. Further, gold money is difficult to counterfeit, because anyone who is familiar with gold can easily distinguish it from cheaper substitutes.

When gold is the real basis for the money of a country, that country is said to be on the *gold standard.* Its basic monetary unit is defined as being equivalent in value to a definite weight of gold. An important feature of the traditional gold standard is *free coinage.* Under free coinage the mint accepts unlimited amounts of gold bullion and returns it to the owners in the form of gold coin if they so desire. Another important feature of the gold standard is redemption of other forms of money in gold coin. Anyone who has paper money or small coins can always exchange these for gold on demand. As a result, paper money and minor coins always have the same value as gold money.[1]

Under the gold standard the weight of gold in the standard monetary unit, such as the dollar, is set by law, and the bullion value of a gold coin is always equal to its money value. Thus, when the United States was on the full gold standard and gold was selling for $20.67 an ounce, a gold dollar contained 23.22 grains of pure gold, or almost one-twentieth of an ounce. This meant that a $20 gold piece contained nearly an ounce of gold.[2] In 1934 the gold content of our dollar was reduced to 13.7 grains of pure gold. Thus, gold bullion became $35 an ounce.

Another metal once favored as a monetary standard is silver. Before the nineteenth century it was much more widely used for this purpose than was gold.

Bimetallism. In the late eighteenth century and in the nineteenth century a number of countries attempted to use both gold and silver as monetary standards. Such a system is known as *bimetallism.* Under it, both metals are coined freely, but the standard monetary unit, like the dollar, is given one weight if of silver and another weight if of gold. The weight of the silver dollar relative to the weight of the gold dollar is

[1] A variation of the gold standard is the gold bullion standard. Under this plan no gold coins would circulate, but the government would either buy or sell unlimited amounts of gold bullion at the same fixed price.

[2] The gold dollar itself was not coined. Gold was coined only in denominations of $2.50, $5, $10, and $20.

known as the *mint ratio*. When the United States established a coinage system in 1792, it provided for a bimetallic standard. Originally the ratio was 15 to 1; later it was changed to 16 to 1. This meant that, since silver was much cheaper than gold, the weight of silver in a silver dollar was 15 or 16 times as great as the weight of gold in a gold dollar.

Unfortunately, bimetallism did not work very well. Although the mint ratio was fixed by law, the *market ratio*, or the relative values of the two metals in the open market, tended to change with changing conditions of supply and demand. If, for example, the demand for silver increased, speculators would not bring silver to the mint to have it coined; rather they would withdraw silver coins from circulation, melt them, and sell the bullion to those willing to pay higher than mint prices. As a result, silver coins would disappear from the monetary system. So also with gold if it increased in value more than silver. Always the metal that was relatively cheap would drive out of circulation the metal that had become more valuable. This tendency of "cheap" money to drive "dear" money out of circulation is known as *Gresham's law*.[3] When one metal had been driven completely out of circulation, the other became the actual monetary standard.

Legally the monetary system of the United States was bimetallic for more than eighty years. However, in the early part of this period the actual standard was silver; later the actual standard was gold; and still later, at the time of the Civil War, it became irredeemable paper greenbacks. These greenbacks continued to be our real standard until 1879, when the Treasury began to redeem them in gold; meanwhile, however, in 1873 Congress had dropped the silver dollar from the list of coins to be minted. This action ended bimetallism legally because it ended the free coinage of silver. Later, the coinage of silver dollars was resumed, but only on a limited basis.

Before the end of the nineteenth century, the difficulty of maintaining a bimetallic standard was recognized almost universally, and bimetallism was everywhere abandoned. Most countries by that time had adopted the gold standard. A few, however, like China and Mexico, still retained the single silver standard.

[3] Sir Thomas Gresham was an official under Queen Elizabeth. He was by no means the discoverer of the principle that bears his name, but he was one of those who early called attention to it.

Paper standards. Gold has two great advantages as a monetary standard: (1) It is the traditional basis for money and is acceptable all over the world.[4] (2) Gold money cannot become worthless because it is always in demand and is available in rather limited quantities.

In spite of these advantages, gold has certain disadvantages which have caused some governments to abandon it or to substitute paper for gold. The disadvantages are (1) the gold available to a particular government often is limited in amount, (2) a gold standard is very expensive to maintain, (3) the outflow or inflow of gold may cause violent changes in prices and this may initiate booms and depressions.

A paper standard can easily be substituted for a gold standard if individuals have become accustomed over the years to the use of paper. However, the amount of paper issued must be limited. The issue of greenbacks by the United States during the Civil War is an example of a paper standard; but in this case the government issued too much paper money and its value depreciated. It was necessary, therefore, to make provision for the redemption of these greenbacks in gold. During the depression of the 1930's, many countries went to a paper standard. In countries where the quantity of paper money issued was limited, this money did not depreciate greatly in value. The experiences of these countries with the paper standard during the depression have caused many economists to insist that the paper standard can be as satisfactory as a gold standard.

Today the United States and most other countries have adopted a form of paper standard, in the sense that no standard gold coins are in circulation and that the paper money which makes up the greater part of the supply of pocket money is not redeemable in gold. However, the tie with gold is by no means completely broken. Gold reserves are still kept, and in the case of the United States they are very large. By this means greater confidence is maintained in our paper money. Moreover, gold is still used to settle foreign trade balances. Gold also plays a large part in stabilizing foreign exchange rates and extending international credits. Nevertheless, it is doubtful whether gold actually exerts much

[4] If all countries were on the gold standard and there were no restrictions on the export or import of gold, foreign exchange rates would be approximately stable. With minor allowances for the cost of shipping coin or bullion, the standard monetary units of any two countries would exchange on the basis of their relative gold weights.

influence on either the quantity or the value of the paper money now in circulation. The value of gold itself seems to be determined largely by monetary policies, especially those of the United States. Any independent influence exerted by gold on the value of money is indirect and is felt only in the long run.

Bank Money

On the basis of the agency that creates it, money may be divided into two classes: *government money* and *bank money.* Government money is important because it furnishes us with our supply of pocket money and also because it serves as the monetary standard. In a sense, however, bank money is even more important, because it furnishes, in the form of demand deposits, by far the greater part of our total money supply. In the United States most payments of any size are made, not with paper or coin, but by drawing checks against demand bank deposits.

Besides deposits there is another form of bank money, the bank note. At one time bank notes, rather than deposits, formed the principal part of our money supply, but over the years their relative importance has gradually declined.

Bank notes. In the earlier days of banking most banking systems allowed ordinary banks to issue noninterest-bearing notes that could be redeemed at the issuing bank at any time. It was soon discovered that, since they were redeemable on demand, people would accept such notes freely as a medium of payment. Thus bank notes became a substitute for metal or paper money issued by the government.

Frequently, however, the banks issuing such notes failed and the notes became worthless. In the United States before the Civil War such failures were frequent. Finally some states refused to permit banks to issue notes unless they first deposited adequate security with a designated state official.

After the passage of the National Banking Act of 1863 the note-issue privilege was restricted to the national banks and was carefully regulated. Finally, with the passage of the Federal Reserve Act in 1913, provision was made for the ultimate retirement of the national bank notes. This ended the note-issue privilege for ordinary banks in the United States. At present, national bank notes are no longer in active circulation, and although some are still outstanding, these are being gradually retired as they come into the hands of the banks.

While the Federal Reserve Act provided for the ultimate retirement of national bank notes, at the same time it also provided for the issue of two new types of bank notes: Federal Reserve notes and Federal Reserve bank notes. Of these two the second type is relatively unimportant.[5] Federal Reserve notes, on the other hand, constitute the greater part of our pocket money.

Federal Reserve notes will be dealt with at some length in the following chapter. At this point it is only necessary to make one observation about them. Though they are issued by the Federal Reserve banks and are called bank notes, they are direct obligations of the Federal government. This means that they are in fact a form of government money. They thus have very little in common with the early bank notes which were based only upon the credit of the bank issuing them.

Bank deposits. Even more important in our present-day monetary system than notes issued by Federal Reserve banks are the deposits of the commercial banks, especially their demand or checking deposits. To the bank, a checking deposit is an obligation to pay a certain amount of paper money or coin on demand. To the depositor it is the right to receive a certain amount of such money.

Often people think of a bank deposit as paper money or coin actually held in the bank for the benefit of the depositor. This, as many of our readers doubtless know, is not the case. The typical bank probably does not hold in its vaults in paper and coin as much as 5 per cent of its total deposits. Even the legal cash reserve that a bank is required to hold is not (if the bank belongs to the Federal Reserve System) paper money or coin. Instead it is merely a deposit in a Federal Reserve bank. If a bank deposit literally represented cash in a bank, then it would be impossible

[5] Federal Reserve bank notes were provided for in the original Federal Reserve Act. They differed from Federal Reserve notes in several respects, including reserve requirements. Originally, they were intended to replace national bank notes as the latter were retired from circulation. Later provision was made for replacing the Federal Reserve bank notes themselves by Federal Reserve notes. When the United States entered the Second World War, most of the Federal Reserve bank notes had been withdrawn from circulation. However, the government had on hand several hundred million dollars' worth of them which had never been issued. As a war measure, to save a quantity of valuable paper and the labor used in engraving, these unissued Federal Reserve bank notes were put in circulation in place of an equivalent amount of Federal Reserve notes. As they wear out, they are again being withdrawn from circulation.

for the total volume of bank deposits to exceed the total volume of cash. Actually, however, the total volume of demand bank deposits held by the public in this country is always several times the total amount of paper money and coin in circulation.

Most economists today classify bank deposits as a form of money. Deposits represent purchasing power just as much as does pocket money. Since the quantity of deposits is much greater, they have even more influence on the prices and production of goods and on the volume of employment than does pocket money. The holder of a bank deposit can of course always turn it into pocket money if he wishes. In order to spend it, however, he does not have to do this. As a rule, all he need do is to write out a check.

The man who receives a check may, of course, cash it; but more often he simply deposits it to his own account. If he deals with the same bank as the man who wrote the check, the bank merely deducts the amount of the check from one account and adds it to another. In such a situation the transaction is a bookkeeping transaction only with no transfer whatever of actual cash. If the receiver of the check deposits it in a different bank, the bank receiving the check for deposit can demand cash over the counter from the bank on which it was drawn. However, most of the claims of banks on one another are canceled out through the clearinghouse so that the balances remaining to be paid are relatively small. In other words, under present methods of doing business the transfer of bank deposits by check does not always mean the transfer of cash. To a large extent bank deposits actually replace pocket money as a medium of exchange. It is estimated that in this country some 85 or 90 per cent of the total volume of payments is now made by checks drawn against demand deposits.

Money in the United States

In 1933 the United States had been on the traditional gold standard, without interruption, for more than fifty years; but following the banking crisis of that year, several important modifications were made in our monetary system. In their final form, these changes were embodied in the Gold Reserve Act of January, 1934. One change was the reduction of the gold content of the dollar by 41 per cent; another was the with-

drawal of all gold coins and gold certificates [6] from circulation and the concentration of the gold reserves of the country in the United States Treasury. Since 1933 none of our money in circulation has been redeemable in gold; we are really, therefore, on a paper standard.

However, as has already been noted, our present paper standard is supported by large gold reserves. It is further tied to gold by the fact that the Treasury, as has been mentioned, now has fixed the price of gold at $35 an ounce. This is at the rate of $1 for about 13.7 grains, which is now the nominal gold weight of the dollar. The Treasury both buys and sells gold at this price. However, within the United States sales of gold are restricted. Since 1933 private hoarding of gold has been illegal. It is now sold only in limited amounts and by special permit for use in industry and the arts.

At present there are three kinds of money in use in the United States: (1) token coins, from the silver dollar down to and including the cent, (2) several varieties of paper money issued either directly by the government or by the Federal Reserve banks acting as agencies of the government, and (3) demand or checking deposits in the commercial banks.

As to coins and paper money, there is little difference between them except in appearance. One is "printed" on metal, the other on paper. Both are token money only.[7] The value of a silver coin is greater than, and is independent of, the value of the metal in it. Coins now made of silver would be worth just as much if made of copper or some other cheap metal. Aside from providing a market for the silver producers, the only arguments for making coins of silver are that they are more attractive and counterfeiting them is more expensive.

As used in the *Federal Reserve Bulletin,* the term "money in circulation" refers to paper money and coin outside the Treasury and the Federal Reserve banks. Table 1 shows the kinds and amounts of such money in circulation, both on Dec. 31, 1939, and on Dec. 31, 1945. By comparing the two sets of figures the reader can readily see the effect of the war in inflating the volume of pocket money.

[6] Gold certificates are still held by the Federal Reserve banks as reserves, but they may not be paid out to member banks or to the public.

[7] Token money as the term is used here is any kind of money, metal, or paper, whose value is greater than, and independent of, the value of the material in it.

Table 1.—Pocket Money in Circulation in the United States

(Millions of dollars)

	1939	1945
Coins:		
Silver dollars	45	139
Subsidiary silver coins	381	832
Nickels and cents	164	307
Paper money:		
Federal Reserve notes	4,912	24,388
Silver certificates and Treasury notes of 1890 *	1,555	1,873
United States notes (greenbacks)	272	316
Federal Reserve bank notes †	24	494
National bank notes	175	117
Gold certificates	69	51
Total	7,597	28,517

Treasury notes of 1890 were called in by the Treasury many years ago, but a small amount of them has never been presented for redemption. Presumably, those notes are lost, destroyed, or held by collectors. For many years the *Federal Reserve Bulletin* listed Treasury notes of 1890 in circulation at $1,000,000. Now any still outstanding are included in the figure for silver certificates.

† Federal Reserve bank notes are explained briefly in footnote on p. 56.

A glance at the table will show that in 1945 Federal Reserve notes constituted more than 85 per cent of all our paper money and coin in circulation. Of the other kinds of paper money, only silver certificates, United States notes, and Federal Reserve bank notes are still in active circulation, and the last named are being replaced by Federal Reserve notes as fast as they wear out. However, considerable amounts of national bank notes and gold certificates are still outstanding. Presumably most of these are being held in hoards in mattresses or safe-deposit boxes. The reader will note that the total volume of pocket money in this country increased by almost four times during the period covered by the Second World War.

As has already been noted, bank deposits are of even greater quantita-

tive importance than pocket money. From December, 1939, to December, 1945, demand bank deposits in the United States in the hands of the public increased from nearly 30 to about 76 billion dollars. In the same period the total of demand and time deposits together increased from nearly 58 to about 149 billion dollars. These figures give the reader some idea of the extent to which money and bank credit were increased as a result of the war.

One may well ask, where did all this new money come from? The answer is that the government created it by the methods it used in financing the war. Had war costs been paid entirely out of taxes and funds borrowed from individuals, there would have been no substantial increase in the amount of money in circulation. This, however, was not the case. To a substantial extent the costs of war were met simply by creating whatever money was needed to pay them. Of course, the government did nothing so crude, simple, and old-fashioned as to print its own notes. Instead, whenever taxes and the sale of bonds to individuals and business firms did not yield sufficient funds, it sold bonds to the banks to make up the difference. The banks bought these bonds by giving the government deposit credits, with the result that whenever the banks as a group increased their holdings of bonds, an equal amount of new bank deposits was created. When the government paid its bills, these deposits were transferred to the general public. Those who received them could then, if they wished, ask for cash; and any new cash demanded was furnished by the Federal Reserve banks in the form of Federal Reserve notes. This process of monetary expansion will be explained in some detail in the following chapter.

Summary

Our study of money and money standards has indicated that money issued by the government, either directly or through the Federal Reserve banks, forms a basis for our monetary system. However, bank money in the form of demand deposits has become the most common means by which exchanges are facilitated.

The gradual shift from a gold to a paper standard has also been noted, as well as the fact that the element of scarcity is a very significant factor in keeping a money system sound. Our study further indicates that, since bank money is so significant, bank issues must be carefully controlled, for banks as

well as governments may create excessive amounts of money unless restricted by law. We noted that there was a great increase in money in circulation during the war, and that this additional money was created by the government in cooperation with the banks.

Because of the significance of banks and bank money, we shall devote the following chapter to this topic before we draw some general conclusions relative to the significance of money to business relationships.

Study Questions

1. Define money.
2. Why is it important for the businessman to understand the principles that relate to money?
3. What is barter? State and illustrate each of the three advantages of trading with money instead of trading by barter.
4. What are the primary or basic functions of money? Name two secondary functions. Why are these latter functions called "secondary"?
5. State four things that are true under the traditional gold standard.
6. Suppose a government provides for free coinage of both gold and silver under a mint ratio.
   a. Will gold and silver necessarily exchange at this ratio in the market? Why or why not?
   b. Will both metals be used as money? Explain.
7. For many years before 1933 the people of this country used both gold and silver coins. Was our monetary system bimetallic? Why or why not?
8. What are the two great advantages of the gold standard over an irredeemable paper standard completely divorced from gold?
9. Why has irredeemable paper money usually depreciated? How can such depreciation be prevented?
10. Why is it often said that the United States is now on a paper money standard? Would it be more accurate to call this a token money standard?
11. State at least two ways in which our money is still tied to gold.
12. Distinguish between government money and bank money.
13. Name the two types of bank money, and explain the nature of each.

14. Which type of bank money now represents the greater part of our money supply? Was this always the case? Explain.

15. Why should Federal Reserve notes be regarded as a form of government money?

16. In what form do member banks of the Federal Reserve System keep their legal "cash" reserves?

17. Are bank deposits usually regarded by economists as a form of money? Why or why not?

18. What are the three types of money now used in the United States?

19. Why is there now little difference, except in appearance, between our paper money and our coins?

20. Which kind of paper money now constitutes more than 85 per cent of all pocket money in circulation?

21. Gold certificates as well as gold were withdrawn from circulation by the Gold Reserve Act of January, 1934. How do you account for the fact that, according to the *Federal Reserve Bulletin*, $51,000,000 worth of gold certificates were still in circulation on Dec. 31, 1945?

22. What was the source of the wartime increase in our money supply?

Exercises

1. Originally, by the Act of 1792, the gold dollar contained 24.75 grains of pure gold. The mint ratio was 15 to 1. How much silver was there in the silver dollar?

2. The gold content of the gold dollar was reduced in 1834, and finally fixed in 1837 at 23.22 grains. The silver dollar was not changed. What was the new mint ratio? (Carry your answer to two decimal places.)

3. An ounce of gold is now the equivalent of $35. How many grains of gold are equivalent to $1? (Carry your answer to two decimal places. A troy ounce, used in weighing precious metals, contains 480 grains.)

4. What percentage of our pocket money consisted of Federal Reserve notes in 1939? What percentage in 1945? Why did Federal Reserve notes increase in quantity so much more than other forms of pocket money?

5. What was the percentage of increase in pocket money from 1939 to 1945? What was the percentage of increase in demand bank deposits during the

same period? Can you give any good reasons to explain the relatively greater increase in pocket money?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 23.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Part I, Chap. 13.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 8.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 11.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 19 and 20.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 18.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 11.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 19.

Spahr, W. E., and Others: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. I, Chaps. 20 and 21.

Taussig, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 19.

6. The Commercial Banking System

Commercial banks. Commercial banks are the ordinary banks that carry checking and savings accounts, make loans, and provide various other services for businessmen and the general public. In this chapter we are interested in them principally because their checking or demand deposits make up the greater part of our supply of money. As we shall see later, these banks do more than merely "carry" deposits. By their lending and investing operations they actually create them.

Bank reserves. Since checking accounts are payable on demand, at first thought one might suppose a bank would always have to hold enough cash to meet such deposits in full. Actually this is not true. Once a bank is established and doing a normal business, it finds that the inflow of cash tends to equal the outflow. It is always making new loans and paying out money against checks presented to it. On the other hand, old loans are always being paid off, and customers are constantly making deposits of cash. Because the inflow of cash tends to equal the outflow, bankers long ago discovered that their cash holdings need be only a fraction of their deposits. The cash that a bank must hold to meet the demands of depositors is known as its *reserve*. The fact that a relatively small reserve can support a much larger amount of deposits explains in part why the total quantity of bank deposits in the country can be several times greater than the total quantity of cash.

In the United States a certain percentage of reserve against deposits is required by law. In addition, every bank must have till money to take care of its ordinary day-to-day transactions. Under the Federal Reserve System all legal reserves of member banks must be deposited with the Federal Reserve bank of the district.

Our Banking System Before and After 1913

It is not possible in a general text on economics to go into the nature and history of our banking system in all its details. Our primary interest is in those characteristics of the system which affect credit policy and the total volume of bank deposits. We are interested in credit policy and deposits because these have a direct effect on the price level and upon business activity in general. For our present purposes it will be especially useful to note certain changes that were made in our banking system in 1913, as a result of the establishment of the Federal Reserve System.

In some respects our commercial banking system has changed very little since 1913. Then, as now, the banks were divided into two groups: state banks, or those chartered by the various states; and national banks, or those chartered by the Federal government. Then, as now, there were many banks of each type; most of them were relatively small and some were not too well managed. It is only fair, however, to point out that the average size of the banks has increased since 1913; and in most cases, it seems probable, the quality of the management has improved. In part this is a result of the failure of many weak banks in the early 1930's; in part it is a result of more adequate bank inspection and supervision.

However, in two important respects our banking system is very different from what it was before 1913. In the first place, before that year the ultimate bank reserves of the country consisted of government paper money and coin held in the vaults of the commercial banks themselves. For practical purposes the total amount of such money at any given time was *inelastic;* that is, there was no way of increasing it to meet unusual needs. If, as sometimes happened, the public demand for cash rose sharply, bank reserves dropped to dangerously low levels. The result was likely to be bank runs, bank failures, a financial crisis, and general business depression. At present there is small danger of the banks getting into trouble merely because of shortage of cash. Any bank which has good assets, and which belongs to the Federal Reserve System, can now borrow additional reserves from a Federal Reserve bank. If it needs cash to pay out over its counter, it can get Federal Reserve notes. The Reserve banks within certain ultimate limits, which will be explained

later, can create these notes in any quantity required. It is said, therefore, that we now have an *elastic* currency.

Our present banking system differs in a second important respect from the earlier one. Before 1913 each bank was largely independent in determining its credit policies. There was no central bank or other coordinating agency. Now, however, the Federal Reserve System exists and exerts very considerable power over the credit policies of the commercial banks as a group.

Federal Reserve System

The Federal Reserve Act of 1913 did not create a single central bank. Instead it provided for dividing the country into a number of districts, and for the creation of a central bank in each district. Actually twelve Federal Reserve districts were established with a Federal Reserve bank in each. Every national bank was required to belong to the system, and state banks were permitted to join if they met certain conditions. The stock of each Federal Reserve bank was to be owned by the member banks of the district, and these banks also were to elect a majority of the board of directors. The Federal Reserve banks are, therefore, not government-owned institutions.

Though the reserve banks are not government-owned, they are subject to a considerable amount of direct government control. The body that now exercises this control is the *Board of Governors of the Federal Reserve System* at Washington, which consists of seven members appointed by the President of the United States. It has direct representation on the board of directors of each Federal Reserve bank because it appoints three of the nine members of each board. It also controls the Federal Reserve System in a number of other ways, some of which we shall have occasion to explain later.

Reserve banks and the member banks. The reserve banks are primarily bankers' banks. Ordinarily they do not deal with the general public but receive deposits from and make loans to other banks only. When member banks are temporarily short of funds, they can borrow from the reserve banks. They must, however, present satisfactory collateral. Most commonly this collateral takes the form of government bonds or commercial paper. By commercial paper we mean the notes or acceptances that the member banks receive when they make short-term loans to businessmen.

Member banks are required by law to keep their entire legal reserve on deposit with the Federal Reserve bank of their district. The legal reserves that member banks must hold against demand deposits vary. Banks are grouped according to the size of the towns in which they are located, and for each group the Board of Governors may determine the reserve ratio within certain limits. For "country" banks the ratio may vary from 7 to 14 per cent; for reserve city banks, from 10 to 20 per cent; and for central reserve city banks (New York and Chicago), from 13 to 26 per cent. The purpose of requiring the member banks to deposit their legal reserves in the Federal Reserve banks is to concentrate the reserves of the banking system. In this way they are made available as the basis for extending credit to any bank that needs help.

When a member bank borrows from a Federal Reserve bank, it ordinarily takes the proceeds of the loan in the form of a credit to its account with the reserve bank. If its purpose in borrowing is simply to maintain its legal reserve ratio, it will leave the borrowed funds on deposit; but if it needs pocket money to pay out to depositors, it will of course draw out cash. The Federal Reserve bank can, as we have seen, create this cash by issuing additional Federal Reserve notes.

Federal Reserve bank reserves. The reserve banks themselves must hold reserves against their deposits and also against Federal Reserve notes. Originally the requirement was 40 per cent against notes and 35 per cent against deposits. The reserve for notes had to be gold, but that for deposits could be held in either gold or any type of lawful (legal tender) money issued directly by the government. When gold was nationalized following the banking crisis of 1933, the Federal Reserve banks were required to deposit their gold with the Treasury in return for gold certificates, and they were also authorized to use these certificates for reserves in place of the metal itself. Later, by an act of June, 1945, the reserve ratios themselves were substantially reduced; and at the same time the use of lawful money as reserve against deposits was eliminated. The present reserve requirement for the Federal Reserve banks is 25 per cent in gold certificates against both deposits and Federal Reserve notes.

The ability of a reserve bank to lend to member banks depends almost entirely upon the amount of surplus reserves that it holds. If its reserve ratio is down to the 25 per cent minimum, it cannot extend deposit credit or issue notes without impairing its legal ratio. On the

other hand, if its reserves are well above the minimum, it is in a position to lend freely. In normal times the reserve banks make it their business to carry substantial excess reserves so that they will be in a position to meet all legitimate demands for credit in times of need.

Federal Reserve notes. By authorizing the issue of Federal Reserve notes, the Act of 1913 created a new type of bank note. This note is an obligation of the Federal Reserve banks but is also a direct obligation of the government. The present provision is that Federal Reserve notes must be secured 100 per cent by collateral in the hands of the Federal Reserve agent.[1] This collateral must consist of not less than 25 per cent in gold certificates and not more than 75 per cent in commercial paper [2] or government bonds. If the reserve banks choose, they may use as much as 100 per cent of gold certificates as the security behind notes.

A note had long been needed that could be issued in increased quantities whenever the banks experienced unusually heavy demands for cash. The procedure for creating additional notes is as follows: The Federal Reserve bank turns over to the Federal Reserve agent not less than 25 per cent in gold certificates and in addition enough government bonds or commercial paper to raise the total security to 100 per cent. It then receives Federal Reserve notes which can be lent to member banks or paid out to them when they draw cash against their accounts. Since a minimum of 25 per cent in gold certificates is needed as reserve, if enough other collateral is available, every dollar in gold certificates is sufficient to permit the issue of four dollars in Federal Reserve notes.

Those who originally drew up the Federal Reserve Act intended that Federal Reserve notes should be elastic in the special sense of reflecting closely the changes in the volume of commercial paper, and hence in the volume of business activity. Their expectations were largely disappointed from the beginning, because experience soon showed that there was no dependable relationship between the volume of commercial paper held by the reserve banks and the demands of the public for cash. To get around this difficulty, since originally the use of government

[1] A Federal Reserve agent is appointed by the Board of Governors as its direct representative with each Federal Reserve bank. He is not only a member of the board of the bank, but he also acts as its chairman. He is not, however, the chief executive officer. To fill this position, the board of each reserve bank elects a president.

[2] The term "commercial paper," as used here, refers to such instruments as drafts, acceptances, and promissory notes growing out of short-term loans to businessmen.

bonds as collateral was not legal, the reserve banks often issued Federal Reserve notes secured principally by gold.[3] However, after the Glass-Steagall Act of 1932 made it possible to substitute government bonds for commercial paper, these bonds came to be used more and more as the principal collateral behind Federal Reserve notes. Today the amount of commercial paper behind the notes is relatively insignificant, and except for gold certificates, they are secured almost entirely by government bonds.[4]

How bank deposits are created. Banks exist primarily to carry deposits and to extend credit. Commercial banks provide businessmen with short-term credit to carry on many of their everyday transactions. For example, to meet seasonal demands for funds merchants will borrow periodically from a bank for short periods. When a bank makes a loan, it may be willing to give the borrower cash at the teller's window. However, it generally prefers to extend to him a deposit credit by entering the amount of the loan in his bank book. By lending in this way the bank avoids any immediate reduction in its cash reserves. At the same time it creates a new bank deposit that previously did not exist. A bank deposit, it will be recalled, is merely the obligation of a bank to pay out money.

A generation or so ago commercial banks employed most of their funds in making short-term loans to businessmen. They still regard such loans with great favor; but in most recent years the demand for them has been limited, and the banks have more and more been obliged to resort to investing in long-term securities, especially government bonds. Commercial loans did rise rapidly in 1946 and early 1947, but bankers were not sure that the demand for them would continue beyond the postwar boom period.

Investments in government securities also bring about the creation of bank deposits. We have already pointed out that during the Second World War bank deposits in the United States more than doubled. Practically all of this increase was the result of the purchase of government bonds by the commercial banks to help finance the war.

The process by which these deposits were created was very simple. The purchase of newly issued government bonds is a way of extending

[3] See "Banking Studies," Federal Reserve System, Washington, D. C., 1941, p. 77.

[4] On Feb. 20, 1946, the collateral held against Federal Reserve notes was as follows: gold certificates, $11,334,000,000; government securities, $14,563,160,000; eligible commercial paper, $391,725,000. See *Federal Reserve Bulletin,* April, 1946.

a long-term loan to the government. When a bank bought, say, $5,000,000 worth of bonds, it was permitted to make payment by merely crediting the United States Treasury with a $5,000,000 deposit. Later the Treasury in paying its bills would draw checks against this deposit and gradually transfer the funds to the public. It should be carefully noted, however, that the new deposit remained in existence and became an addition to the amount of deposits already held by the public. Furthermore, most of the new deposit money created during the war will stay in circulation indefinitely. The only thing that would be likely to bring much reduction in the combined volume of deposits and pocket money would be a substantial reduction in the government debt.

We are now prepared to show how bank deposits can be expanded. We shall find that if a bank receives new cash from outside the banking system, for example, cash representing newly mined gold or imported gold, the individual bank will not be able to expand loans by much more than that part of this cash which represents surplus reserve. Even if it makes loans in the form of deposit credits, it will be able to retain comparatively little of the deposits so created. On the other hand, we shall find that the whole system of banks can create new loans and deposits equal to several times the amount of any new cash which is obtained.

Let us suppose that a certain bank has extended loans to the limit of its ability, and that it then receives a deposit of $1,000 in cash. Let us suppose further that this cash represents gold imported from abroad. The balance sheet of the bank would then contain the following:

Assets		*Liabilities*	
Cash	$1,000	Deposits	$1,000

However, if we assume that this bank is required to carry a legal reserve of 20 per cent, it needs only $200 of the new cash as reserve against the new deposit. This $200 it must deposit with the Federal Reserve bank of its district. The other $800 is available for expanding loans.

Now let us suppose that a customer comes to the bank and desires to borrow $800 on his note. The bank might hand over to him $800 in cash. This is not likely, however, for two reasons: (1) the customer will probably prefer to pay his bills by check; (2) the bank would like to keep the cash as reserve as long as possible. Consequently, when the bank makes the loan, it credits the depositor's account with $800.

We should observe here, however, that whether the bank credits the customer's account or gives him cash over the counter does not make so much difference as is often supposed. People who borrow money generally borrow it to use. If the bank lent the customer $800 in cash, it would lose the cash immediately; if it credits the customer's account, he will very soon draw checks against it for most of the $800, and most of these checks will be presented at other banks by the people who get them. The borrower, to be sure, may not check out *all* the $800 (he may have agreed with the bank to leave a certain amount, say 20 per cent, on his account), and some of the checks he draws may come back to his own bank. For simplicity of exposition, however, let us assume that he checks out the whole $800, and that all the checks drawn are presented at other banks for payment. Our bank will then owe $800 to the other banks through the clearinghouse, $800 *that it would not have owed had it not made this particular loan;* and when it has met this obligation, it will have paid out $800 in cash and its balance sheet will show the following changes:

Assets		*Liabilities*	
Cash	$200	Deposits	$1,000
Loans	800		

Our original bank, as a direct result of receiving the $1,000 cash, still has an increase in deposits of $1,000, and it has expanded loans by $800. It is not now in a position to expand either loans or deposits further. For the banking system as a whole, however, the original $1,000 in new cash has by no means reached the limit of its power to create loans and deposits.

If we assume that the people who received our borrower's checks for $800 deposited them in other banks *instead of asking for cash,* these other banks will have $800 in new deposits. They will also have $800 new cash, received from our bank through the clearinghouse. There is now $1,800 on deposit in the whole banking system; that is, the $1,000 deposited in the first bank plus the $800 deposited in other banks. The banks that have just increased their deposits and cash by $800 must deposit 20 per cent of the cash, or $160, with the Federal Reserve bank as legal reserve against the new deposits. However, this leaves them with $640 in new cash with which to make $640 in new loans. When in turn the proceeds of these loans are checked out and deposited in other banks, $640 *more*

of new deposits will have been created in the system; and this process can be repeated again and again. If and when it reaches its limit, the original $1,000 in new cash will have enabled the banks as a group to expand deposits to $5,000. When that point is reached, all the $1,000 in cash will be on deposit with the Federal Reserve banks to provide the necessary 20 per cent reserve against the $5,000 of new deposits.

However, $5,000 is not the maximum amount of demand deposits that our banking system can create on the basis of $1,000 new cash, if that cash is acquired in the form of gold. When the member banks deposit the gold in the Federal Reserve banks,[5] the Federal Reserve banks are required to turn it over to the Treasury in return for gold certificates. The certificates constitute reserves for the reserve banks themselves, but these banks are required by law to keep reserves of only 25 per cent against their own deposits. Against $1,000 in deposits they need hold only $250 in gold certificates.

The Federal Reserve bank statement then reads, so far as this one transaction is concerned, as follows:

Cash (gold certificates)	$1,000	Deposits	$1,000

We may state the same situation in another way:

Necessary reserves	$250	Deposits	$1,000
Excess reserves	750		

The Federal Reserve bank now has $750 in excess reserves in the form of gold certificates. These can be the basis of further credit expansion in either of two ways. First, the reserve bank can use them as the 25 per cent reserve it must hold against notes, and with the addition of 75 per cent of commercial paper or government bonds, they can be used to issue $3,000 in Federal Reserve notes ($750 divided by 25 per cent). Second, member banks can borrow from the Federal Reserve banks to increase legal reserves, that is, their deposits in the reserve banks.

The process of bank borrowing is as follows: If a member bank wishes to borrow from the Federal Reserve bank to increase the amount of actual cash for its everyday business, the Federal Reserve bank will rediscount the commercial paper that is presented to it. In such a case under the assumptions mentioned above, the Federal Reserve bank can

[5] They are required by law either to deposit it or to exchange it for some other form of money.

rediscount a sufficient amount of commercial paper to equal $3,000 if the borrowing bank demands cash, and this amount will be paid to the member banks in Federal Reserve notes.

Suppose, however, the bank wishes to exercise its other alternative, namely, to increase its deposits with the Federal Reserve bank. In this case the Federal Reserve bank can lend $3,000 as before, but in the form of a deposit credit rather than Federal Reserve notes. The member bank can then use the deposit so created as a reserve for further extension of its own deposit credit to businessmen.

Suppose the member bank should extend new loans in the form of deposit credit, equal in amount to the new reserve. When these loan-created deposits were checked out and the checks deposited in other banks, these other banks would have new deposits of $3,000. If we assume as before an average legal reserve ratio of 20 per cent for member banks, eventually, by the process already described, the banks as a whole could create additional deposits equal to five times $3,000, or $15,000. Thus, with an original increment of $1,000 in new cash, the banking system may theoretically expand its demand deposits to $20,000; that is, before the member banks have exhausted their own reserves, they can expand deposits to $5,000, and then on the strength of the deposit of their $1,000 of reserves in the Federal Reserve banks, these latter banks can create deposits in favor of the member banks of $3,000, which in turn they can use to expand their own deposits to the extent of $15,000 more. In other words, if we assume an average legal reserve ratio of 20 per cent, member banks by depositing their reserves with the Federal Reserve bank can expand their own deposits by five times their reserves. Furthermore, if they borrow from the Federal Reserve bank, they can increase their reserves to four times the actual cash that they deposit with the Federal Reserve bank.

Perhaps Table 2 will help students to understand the process by which a system of banks can expand an original cash deposit of $1,000 in one bank into an amount several times as great for the system as a whole. Again we assume a 20 per cent reserve ratio. Further, to simplify the situation, we shall make several other assumptions: (1) the banks make no use of Federal Reserve credit, (2) each bank is able and willing to lend the maximum amount which the 20 per cent reserve requirement permits, (3) the entire amount that the first bank lends is checked out and deposited in a second bank, what the second bank lends is deposited in a third bank, and so on. If one examines the table, it is obvious that

if this process were carried far enough, the sums involved would become very small, and approximately $5,000 in deposits would have been created by the original $1,000 in cash. In our illustration we have used ten banks just to show the trend.

Table 2

	Additional deposits (100%)	*Additional loans* (80%)	*Additional reserves* (20%)
First bank	$1,000.00	$ 800.00	$ 200.00
Second bank	800.00	640.00	160.00
Third bank	640.00	512.00	128.00
Fourth bank	512.00	409.60	102.40
Fifth bank	409.60	327.68	81.92
Sixth bank	327.68	262.14	65.53
Seventh bank	262.14	209.71	52.42
Eighth bank	209.71	167.77	41.94
Ninth bank	167.77	134.21	33.55
Tenth bank	134.21	107.37	26.84
Totals, first 10 banks	$4,463.12	$3,570.50	$ 892.62
Other banks in turn	536.88	429.50	107.38
Total, all banks	$5,000.00	$4,000.00	$1,000.00

A formula will permit the student to compute quickly the expansion that theoretically can take place. Let D equal the amount of deposits after the expansion takes place, C the amount of new money deposited in the banking system, and R the reserve ratio. Then

$$D = \frac{C}{R}$$

Thus, if the new money deposited is $1,000 and the reserve ratio is 20 per cent, the resulting total volume of deposits is

$$D = \frac{1{,}000}{20\%} \quad \text{or} \quad \frac{20}{100} D = 1{,}000$$

$$20D = 100{,}000$$

$$D = \$5{,}000$$

The amount of expansion in loans is given by the formula

$$L = \frac{C - CR}{R}$$

where C and R have the same meaning as in the deposit expansion formula. Thus,

$$L = \frac{\$1{,}000 - (\$1{,}000 \times 20\%)}{20\%}$$

$$L = \frac{1{,}000 - 200}{20\%}$$

$$L = \frac{800}{20\%} = \$4{,}000$$

We should note one important qualification of the above explanation of bank-deposit expansion. Actually, our illustrations considerably exaggerate the extent to which the banking system can expand on the basis of a given amount of surplus reserves. For the sake of simplicity we neglected to take into account the fact that, as people hold larger average bank deposits, they are almost certain to draw out and keep out larger amounts of paper money and coin. This fact, however, does not affect the power of the Federal Reserve banks to lend to the member banks, but it does reduce the power of the member banks as a group to expand their own deposits on the basis of such loans.

Let us suppose, for example, that the Federal Reserve banks have $1,000,000 of excess reserve in gold certificates. On the basis of this reserve they can lend member banks $4,000,000 if the latter have adequate commercial paper or government bonds for security. If the member banks take the loans in the form of deposits in the reserve banks, *and if as a group they can continue to hold them in that form,* they can then expand their own loans and deposits by as much as $20,000,000. However, as soon as bank deposits begin to expand, there will be an increased demand by the public for pocket money. To meet such a demand the member banks must then draw out some of their Federal Reserve deposits in the form of Federal Reserve notes; and this of course reduces their new reserves of $4,000,000 and also their power to expand loans and deposits further.

It is impossible to determine in advance just how much a given increase in bank deposits will increase the demand for Federal Reserve

notes for hand-to-hand circulation. There is no question, however, that an increasing demand for cash is an important factor in limiting the maximum amount by which the banking system can expand loans and deposits. In the case assumed, this factor might well reduce the possible expansion 50 per cent or more; that is, instead of the banking system being able to expand loans and deposits by $20,000,000 on the basis of $1,000,000 in excess reserves in the Federal Reserve banks, it might actually be able to expand them by no more than $10,000,000.

Summary

The purpose of this chapter has been to describe the commercial banking system and the way in which it creates both bank deposits and pocket money. Emphasis was placed on the Federal Reserve System, the provisions that give elasticity to Federal Reserve notes, and the relations of member banks to the Federal Reserve banks. Finally, the process by which the member banks as a group can expand their deposits by increasing their loans and investments was explained in some detail.

Study Questions

1. What are commercial banks? Are there any other kinds of banks?
2. What is the principal reason for introducing a chapter on the commercial banking system at this point in our text?
3. What is meant by *bank reserves*? Why is it unnecessary for reserves to be as large as deposits?
4. The text emphasizes two changes that the Federal Reserve Act of 1913 brought about in our banking system. Explain each of them.
5. Describe the general structure of the Federal Reserve System.
6. What are the legal reserve requirements for member banks?
7. What are the present legal reserve requirements for Federal Reserve banks?
8. Why do the Federal Reserve banks ordinarily carry reserves considerably above the legal minimum?
9. What great need was met by providing for the issue of Federal Reserve notes?

10. What must the reserve banks do in order to create additional Federal Reserve notes?

11. The authors of the Federal Reserve Act intended Federal Reserve notes to be elastic in a special sense, but their expectations were disappointed. Explain.

12. Aside from gold certificates, what is now the principal collateral behind Federal Reserve notes?

13. Explain simply how a bank can create a new deposit (*a*) by making a loan to a businessman and (*b*) by buying government bonds.

14. Why is it that a single bank with $1,000 of excess reserve cannot safely lend much more than this amount, even if it makes the loan in the form of a deposit credit?

Exercises

1. Suppose that a hoard of $10,000 in gold certificates is discovered and deposited in a member bank. Assume that the legal reserve ratio against deposits is 20 per cent.

a. On the basis of this deposit and without borrowing from a Federal Reserve bank, how much can this member bank safely lend in the form of deposit credits?

b. Again assuming no borrowing from a Federal Reserve bank, how much will the member banks as a group be able to extend their loans and deposits? (Assume an active demand for loans.)

c. The bank receiving the $10,000 in gold certificates is required by law to deposit them in a Federal Reserve bank. How much of the gold certificates must the Federal Reserve bank keep as reserve against this deposit?

d. How much will it be able to increase its loans to member banks in the form of deposit credits?

e. How much can it increase its loans in the form of Federal Reserve notes?

2. If the legal reserve ratio for member banks is 20 per cent, it is possible under certain circumstances for every dollar in surplus reserves (gold certificates) held by a Federal Reserve bank to support an increase of $20 in the deposits of the member banks as a group.

a. Show that this is theoretically possible.

b. Would it be possible if, as bank deposits increased, the public withdrew more Federal Reserve notes from the banks in order to carry larger amounts of pocket money?

c. If the member banks were required to maintain a 30 per cent instead of a 20 per cent reserve, how much deposit credit could the banking system create for every dollar of surplus reserve in a reserve bank? (Assume that the public does not increase its holdings of pocket money.)

3. A member bank that is required to maintain a 14 per cent legal reserve has customer deposits of $50,000.

a. How much reserve must be maintained with the Federal Reserve bank?

b. How much reserve must the Federal Reserve bank maintain on the basis of this deposit by the member bank?

4. In our analysis we have assumed, for the sake of simplicity, that, when people borrow from a bank and receive the loan in the form of a deposit credit, they at once check out the entire amount of the deposit so created. This is not strictly true. A bank that is constantly creating deposits by lending will always have some of these deposits on its books. Such loan-created deposits are called *derivative deposits,* while those created by the receipt of cash are called *primary deposits.*

Suppose a bank must keep a reserve ratio of 20 per cent. If all loan-created deposits are checked out at once, how much can it expand loans on the basis of new cash deposits of $100,000? On the basis of these same cash deposits could it expand loans by a greater amount if its derivative deposits were always about 25 per cent of its loans outstanding? Why or why not?

5. On July 31, 1946, the Federal Reserve banks held $18,105,481,000 of gold certificates. Federal Reserve notes outstanding were $24,244,229,000; total reserve bank deposits were $17,905,819,000.

a. What was the actual ratio of reserves to notes and deposits combined?

b. How much in gold certificates did the reserve banks hold over and above what was necessary to meet their legal reserve requirements?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chaps. 25 and 26.

Burgess, W. R.: "The Reserve Banks and the Money Market," Harper & Brothers, New York, 1936, Chaps. 1-7, 16-19.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 9.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chaps. 21-23.

FEDERAL RESERVE SYSTEM, BOARD OF GOVERNORS: "The Federal Reserve System, Its Purposes and Functions," Washington, D. C., 1939.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 21.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 18.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chaps. 12 and 13.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chap. 20.

MINTS, LLOYD W.: "A History of Banking Theory," University of Chicago Press, Chicago, 1945.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. I, Chaps. 22 and 23.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chaps. 22-25, 27.

7. The Purchasing Power of Money

Thus far we have been concerned with money as a medium of exchange and a measure of value, and we have assumed that its principal effect on our economy is to facilitate exchange. We can easily determine, however, that money is not a very good measure of value from one year to another, and that at times changes in its value have seriously disrupted business relationships.

A change in the value of money means an increase or decrease in its power to purchase goods. When such a change occurs, money becomes something more than merely a medium of exchange or a measuring rod for the value of goods. It becomes an active economic force which brings gains to some and losses to others, and which exerts important effects on the rate of business activity.

Changes in the value of money create two problems: (1) How shall we measure such changes? That is, how can we express the relationship between the value of money and the prices of commodities? (2) Even more important, just what effects do changes in the value of money have upon the economic system, and how do these effects come about?

The value of money. At this point we should recall the definitions of value and price which already have been given. Value was defined as the power of one good to exchange itself for another. Price is the value of goods in terms of money. If goods are valued in terms of money and if the value of money varies, how can its value be measured? Obviously, such value cannot be determined by comparing money with some one good, for this good may vary in value; consequently, the relationship established will not be significant. Moreover, it will do no good to try to value money in terms of itself. The only method, then, that can be

employed to determine the value of money is to measure its purchasing power in terms of economic goods in general. There is, however, no entirely satisfactory way of measuring the *absolute* value of money, that is, of measuring the physical quantity of *goods in general* that money will buy. However, insofar as we can get adequate price statistics, it is not difficult to measure *changes* in the value of money. This is accomplished by comparing the average prices of the goods and services selected as a sample in one period with those in another.

Suppose that in a given year $1,000 will buy all the goods and services that a certain family is accustomed to consume. Suppose, further, that in some later year it requires $2,000 to buy the same goods in the same quantities. In this situation it is clear that (1) average prices, or the general price level as indicated by our sample, must have doubled; and (2) the value of money must have fallen by one-half, since it takes twice as much money as before to buy the same goods. We see, then, that the ratio of prices to the value of money is an inverse one. In other words, when prices go up, the value of money goes down, and vice versa.

How the Value of Money Is Determined

Before attempting to analyze the forces that determine the value of money, it is necessary to fix in mind certain basic relationships between money and goods. It is clear that, if all trade is carried on in terms of money, the money spent in a given period of time must be equal to the money value of the goods and services exchanged. Suppose that in a certain community all money is spent once and only once a year, and that in a certain year there was $1,000,000 in circulation. Suppose, further, that 1,000,000 units of goods were exchanged. Then the total value of the goods exchanged must have been $1,000,000 and the average value of each unit $1. Suppose that the next year the amount of money in circulation increased to $2,000,000 while the amount of goods remained constant. It follows that the money value of all goods exchanged must have risen to $2,000,000 and the average value of each unit to $2. In other words, if the total amount of money spent increases, while the quantity of goods exchanged remains constant, the price level must rise. The contrary proposition is equally true. If the amount of money spent decreases, while the quantity of goods stays the same, the price level must fall.

Equation of exchange. In the paragraph above, for the sake of simplicity, we assumed not only a constant quantity of goods, but also that each dollar was used only once in a given year. Actually, however, the same dollar may be used over again a number of times. If on an average each dollar is used five times in a year, it is clear that at any given price level it can exchange five times the quantity of goods that it could if used only once. For example, the $1,000,000 mentioned above could do as much work as $5,000,000 used only once. We call the average number of times that a dollar is used each year the *velocity* of money.

If we include velocity, we find that there are four major factors involved in the relationship between money and prices: (1) the price level itself, (2) the physical quantity of goods and services exchanged for money in a given period, (3) the average amount of money in circulation, including both pocket money and demand bank deposits, and (4) the velocity, or average rate of turnover of the money. It is possible to express the relationships among these four factors in a simple mathematical formula known as the *equation of exchange.*

In order to simplify our exposition, and also because we have not yet explained just how the general price level is measured, let us assume that all goods and services that are exchanged can be divided into a large number of like units. On this assumption, the price level will be represented by the price of our imaginary goods unit.[1]

We are now ready to construct the equation of exchange. We shall let

P = the price of a goods unit, or the price level
T = the number of goods units, or the physical volume of trade
M = the quantity of money in circulation, including both pocket money and demand deposits
V = the velocity or average rate of turnover of money.

The equation of exchange can now be expressed in the form:

$$PT = MV$$

This simply means that the price of each goods unit, multiplied by the number of units exchanged, equals the money in circulation, multiplied

[1] The only way in which it is possible to conceive of dividing the total physical volume of trade into a number of like units is to think of a composite commodity. Each unit of this commodity would be made up of a very small amount of every material good and service entering into trade.

by the average number of times each unit of money is used. The equation, of course, refers to a given economic system and a given period of time. *PT* (price times units of goods traded) is one way of expressing the total money value of goods sold. *MV* (money in circulation multiplied by the average times each unit is spent) is another way of expressing the same thing. Therefore, the equation of exchange is often called a truism, since in effect it simply says that *the money value of goods sold equals the money value of goods sold.*

The equation of exchange is a truism only on the assumption that all goods sold are paid for with money. It does not take account of barter transactions; nor of credit transactions, unless the latter are settled by cash or check before the end of the period to which the equation refers. The fact that the equation of exchange does not take account of barter is not very important, since there is relatively little barter in a modern economy. The fact that it does not take account of credit transactions is a more serious limitation. It is not, however, so important a limitation as it might at first seem. Credit transactions, if the credit is good, are ultimately settled by cash or check; and although some transactions made this year will not be paid for until next, these tend to be offset by transactions entered into last year and paid for this year.

Since one side of the equation of exchange must always equal the other, if any factor in the equation changes, this variation must be offset by a change in one or more of the other factors. The equation, however, tells us nothing as to which of the other factors will change or how much. We only know that the changes that occur must be such as to keep the two sides of the equation equal.

Let us suppose, for example, that $P = \$10$, $T = 500{,}000$ goods units, $M = \$1{,}000{,}000$, and $V = 5$. Then $PT = MV$ becomes $\$10 \times 500{,}000 = \$1{,}000{,}000 \times 5$, or $\$5{,}000{,}000 = \$5{,}000{,}000$. Now let us suppose that the government prints more money, so that M increases to $\$2{,}000{,}000$. Some change must now take place in the other factors. However, the equation itself does not tell us which factors will change. It may be that P will become \$20, T 1,000,000 units, or V 2½. Any one of these changes would satisfy the equation. Another possibility is that some change will take place in all these factors. To make any valid judgment as to what is likely to happen as the result of a change in the quantity of money, we must go behind the equation itself and study in the real economic world the factors represented in the equation.

The usefulness of the equation of exchange results simply from the fact that it groups together in a convenient way certain closely connected economic factors and helps us to study the relationships among these factors without becoming hopelessly confused.

Sometimes the equation of exchange is expressed in terms of six factors. In this case M stands for paper and coin only, while demand bank deposits are represented by M'. V is the velocity of paper and coin, V' the velocity of bank deposits. The equation can then be written

$$PT = MV + M'V'$$

This means that the average price of a goods unit times the number of units exchanged equals the amount of pocket money in circulation times its velocity, plus the amount of demand bank deposits times their velocity.

This formulation of the equation of exchange was much more significant in the past than it is today. Under the old gold standard, and especially before the establishment of the Federal Reserve System, the volume of demand deposits (M') which the banks could create depended largely on the amounts of cash (M) available for general circulation and for bank reserves. It seemed desirable, therefore, to separate the key factor M from the dependent factor M'. However, it is no longer true that the volume of bank deposits is controlled by the volume of paper money and coin. The reserves of member banks now consist, not of coin and government paper, but of deposits in the Federal Reserve banks; and the amount of demand deposits that the member banks can create depends very largely on Federal Reserve policy and government fiscal policy.

Under present conditions, therefore, once a given amount of bank deposits is created, the amount of pocket money in circulation depends almost wholly on the wishes of the public. If people choose to withdraw a large portion of their deposits in cash, the Federal Reserve banks create the necessary cash in the form of Federal Reserve notes, and as a result M increases while M' decreases. On the other hand, if people prefer deposits and take their cash to the banks, Federal Reserve notes are withdrawn from circulation. M' then increases and M decreases.

Since M and M' both represent immediate purchasing power in terms of money and since people can hold their money in either form they

choose, for our purposes there is not much reason for drawing a distinction between them in the equation of exchange.

The quantity theory. The quantity theory in its various forms is the oldest and most influential of the theories that attempt to explain how the value of money is determined. According to Professor Chandler, crude statements of it are found at least as far back as Roman times.[2] Not, however, until close to the end of the eighteenth century did it receive clear formulations somewhat resembling those held by quantity theorists today.

In ancient times and in the Middle Ages, money meant specie, principally gold and silver coins. In its early forms the quantity theory was the belief that the price level or the value of money depended on the relation between the amount of specie in circulation and the amount of goods available for purchase. Frequently, however, the amount of goods was assumed to be constant. Changes in the amount of money then became the only thing that mattered. This reduced the quantity theory to its crudest form, namely, the belief that the value of money varied inversely with its quantity. This meant, for example, that if the amount of money in circulation doubled, the value of each unit of specie would automatically be cut in half; if the amount of money was cut in half, the value of each unit would double. As time went on, however, the quantity theory was refined. By the end of the eighteenth century, Hume and others had submitted formulations that took into account all four factors represented in what we now call the equation of exchange.

In its modern form the quantity theory is still an explanation of the relation between the quantity of money and the price level, but it is an explanation that makes full allowance for the possible effect of changes in velocity and in the volume of trade. As Chandler puts it, "No competent economist since the time of Hume has believed that there is any fixed and automatic relationship between the quantity and the value of money." [3] All have recognized that changes in the quantity of money affect the price level only by affecting total money expenditures and the money demand for goods. They have also recognized that changes in money expenditures and the price level are not necessarily proportional to changes in the quantity of money. This is partly because the physical

[2] CHANDLER, LESTER V., "An Introduction to Monetary Theory," Harper & Brothers, New York, 1940, p. 21.

[3] *Ibid.*, p. 22.

quantities of goods available in the market may change, but it is even more because of possible changes in the velocity of the circulation of money.

Let us examine by means of an illustration the process by which a change in the amount of money affects the price level. We will suppose that the government, to finance relief payments or war expenditures, sells great quantities of bonds to the banks. As a result, sufficient bank deposits are created to double the amount of money (M) in circulation.

At first the new money will be held by the government in the form of deposits in various banks. Gradually, however, as the government pays its expenses these deposits will be transferred to the public and will be widely scattered. People then, on an average, will hold twice as much money as they did before. Human nature being what it is, they will then increase their expenditures. However, it is unlikely in the short run, as we know from past experience, that their expenditures will increase in proportion to the increase in their stocks of money. Therefore, since expenditures are not increased to the full amount of the increase in money, the average dollar will not be used quite so often as before, and this means that velocity (V) will decline somewhat. Nevertheless, since there is some increase in total expenditures, the money demand for goods and the price level (P) will rise. However, prices will not double. This is partly because expenditures have not doubled and partly because, as prices begin to rise in response to an increased demand for goods, production is stimulated. There is then an increase in the physical volume of trade (T); and this tends to check the price rise.

We see, then, that the net result of the doubling of M in our illustration above was to cause some variations in all three of the other members of the equation of exchange. These variations, taken together, were of the kind and amount necessary to keep the two sides of the equation equal. On the right-hand side M doubled but V decreased somewhat, so that although MV was greater, it did not double. On the left-hand side of the equation both P and T increased, but only enough between them to keep PT equal to MV.

Few modern economists would deny the general validity of the above analysis as applied to the short run, and in this case the short run might cover several years. Many, however, would insist that in the long run, if the amount of money in circulation is changed, the principal effect will be on the price level.

In support of this point of view, it is argued that in the long run the physical volume of goods, and the velocity of circulation of money, are largely independent of the quantity of money. They may be affected temporarily by changes in the quantity of money, but eventually they tend to return to their normal level. This normal level in the case of velocity is determined by established methods of doing business and by such factors as transportation, communication, and other facilities for making payments. In the case of the physical quantity of goods, it is maintained that the normal level is determined by the stage of technological development and by available supplies of labor, capital goods, and natural resources. Since, then, both the velocity of money and the physical quantity of goods are in the long run determined independently by other factors, it follows, according to this point of view, that the ultimate effect of a change in the quantity of money will be felt almost entirely in a change in the price level.

Whether or not we accept the above argument completely, we are safe in assuming that there is a significant relationship between the quantity of money and its value. In the long run the value of money, like that of anything else, tends to decrease as its quantity increases.

Sometimes, however, we are not especially interested in the long run. As one economist has said, in the long run we are all dead. Many of our most serious economic problems, like high unemployment in times of depression, grow out of short-run maladjustments.

In the short run, as has been pointed out, the exact relationship between the quantity of money and the price level is rather uncertain because considerable changes can take place in velocity and the physical volume of trade. Further, not all economists would agree as to the usual sequence of cause and effect. In the illustration given above, the increase in the quantity of money was brought about by government fiscal policy. Any change in prices must therefore have been a result, not a cause, of the increased money supply. In some cases, however, it is conceivable that a rise in prices might precede and then bring about an increase in the supply of money.

Suppose, for example, that the country is nearing the end of a period of depression. Prices and wages are low, but they refuse to drop further. Gradually more and more businessmen decide that it is a good time, because costs are low, to expand plant and equipment with an eye to future needs. At first they are able to finance this expansion out of large

idle bank balances. However, the new demand for labor and materials, and the increased expenditures that workers can now make, will tend to raise prices. This increase in prices will encourage businessmen to expand further and to borrow funds from the banks; and the expansion of bank loans will then create new money in the form of bank deposits. In this case it is clear that a price rise stimulated by business expansion was one of the *causes* of an increase in the quantity of money.

However, even in a case like the one above, to debate whether the rise in prices *caused* the increase in money or vice versa is a good deal like arguing which came first, the hen or the egg. In periods of business expansion prices and bank deposits *both* rise, and there is little doubt that there is a mutual interaction of cause and effect. A rise in prices stimulates the expansion of bank loans and deposits, and the expansion of deposits gives further impetus to the rise in prices. Whether this process of expansion starts with prices or with money need not concern us too much. Of one thing we can be pretty sure. Even in the short run a rise in the price level cannot go very far unless there is some way of creating additional money to support it.

Enough has been said relative to the relationship of money and prices to make clear the fact that changes in the quantity of money cause changes in price relationships. These changes must be taken into account by the businessman and are one of the factors over which he may be able to exert little control. It is of the utmost importance, therefore, not only that he understand the implications of the quantity theory, but that he use the measures that are available to him to determine what effect changes in the quantity of money are having on the prices of the products that he buys and sells. It is, therefore, necessary to explain how the businessman can measure changes in the price level.

Measuring Changes in the Price Level

Since the value of money is its purchasing power, its value changes whenever the average price of goods rises or falls. Consequently we may compare the value of money at different times by determining what changes have taken place in the price level. Shortly we shall explain the relation between the price level and the purchasing power of money. Meanwhile we shall proceed to a study of the methods used to measure

changes in the price level. This information, in turn, will enable us to measure changes in the purchasing power of money.

Index numbers. Price index numbers are percentages that express changes in the price level from one time to another. Thus, if the index in one year is 85 and in another it is 110, prices have increased between these two periods. However, it is obvious that the numbers 85 and 110 must be related to some common base, for otherwise how were the 85 and 110 obtained? The method used is to express the price level of a certain base year as 100 and then to compare the price level in other years with this base. The index 85, then, means that the price level in that year was only 85 per cent of the price level in the base year; whereas an index of 110 means that prices were higher than those of the base year by 10 per cent.

Index numbers may be used for other purposes than to express changes in the price level. They may, for example, indicate the number of carloadings, the amount of construction, or the general level of industrial activity. In this discussion, however, we are interested only in the kind of index number that shows changes in the general price level.

Index numbers of the general price level may be constructed by a variety of methods; but to simplify our exposition, we shall consider only one, namely, that which employs *weighted price aggregates.* This is the method adopted by the Bureau of Labor Statistics, whose price indexes are more widely used than any others.

Method of weighted aggregates. To employ this method we must have two kinds of information: (1) the average price of each commodity in each year for which we wish to find an index and (2) the quantity of each commodity sold in some one selected or base year. For every year we then multiply the price of each commodity by the quantity sold in the base year, and add the products so obtained. This gives us the weighted price aggregate, or sum, for each year. The quantities sold we call the *weights.* The effect of multiplying the prices of commodities by the quantities sold is to make commodities count in the price aggregate somewhat in proportion to their economic importance. However, we must use the same quantities for every year, because we want the aggregates to show only changes in price. Having selected one year as the base, we call the price index for that year 100. To obtain the price indexes for the other years, all that is necessary is to express their

weighted price aggregates as percentages of the weighted price aggregate for the base year.

Table 3 illustrates the calculation of index numbers by the use of weighted price aggregates. Ideally, in calculating an index number, all commodities in the market should be included. The Bureau of Labor Statistics index is based on over 800. However, we are interested only in the method, and to simplify our problem we have reduced the number of commodities to five.

Table 3.–Index Numbers of Prices–Weighted Aggregates Method

Commodity	*Price per unit in the base year* 1938	*Quantity sold in base year* 1938 (000 *omitted*)	*Total value commodities in base year* 1938	*Price per unit in* 1942	*Total value commodities in* 1942
Wheat per bu.	$0.70	400,000	$280,000	$1.00	$400,000
Butter per lb.	0.35	70,000	24,500	0.40	28,000
Coal per ton.	2.00	20,000	40,000	2.25	45,000
Cotton per lb.	0.15	800,000	120,000	0.20	160,000
Beef per lb.	0.25	600,000	150,000	0.30	180,000
Weighted aggregate			$614,500		$813,000
Divide by base year			614,500		614,500
Index number			100		132

The data presented in the table show that the price level increased by 32 per cent between 1938 and 1942 and that the prices of all commodities used increased. This may not always be the case. In fact, when individual prices are compared from year to year, some may be found to increase and others to decrease. Moreover, even when all prices increase, the rate of increase for some commodities is much greater than for others. In the example, for instance, wheat increased by 43 per cent, butter by only 14 per cent, coal by 12½ per cent, and so on. Since we are interested in the price level changes rather than in the changes in any particular commodity, for our purposes the index 132 is significant for the year 1942.

The rise in the price level indicates that commodities rose in value

when compared with money; that is, in this period the value of money declined, for it required only $614,500 in 1938 to purchase a quantity of goods sold for $813,000 in 1942.

Purchasing power of money. We are concerned with the problem of measuring changes in the purchasing power of money, but the index numbers show changes only in the general price level. How then can index numbers be translated into purchasing power? Since the value of money varies inversely with prices, the purchasing power of money declines as prices rise. Stated more precisely, the purchasing power of money is the reciprocal of the price index. Thus, if the index is 100 per cent or 100/100 the reciprocal is 100/100 or 100 per cent. However, if the index is 125 or 125/100 the purchasing power is 100/125 or 80 per cent. The Bureau of Labor Statistics index of prices shows that, based on 1926, the average wholesale commodity price index in 1913 was 69.8 and the purchasing power of money 143.3. In 1920 the price index was 154.4 and the purchasing power of money 64.8; in 1939 these numbers were 77.1 and 129.7, respectively; in late 1947 they were approximately 158 and 63.3.

Effects of Changes in the Value of Money

Since changes in prices mean changes in the purchasing power of the dollar, persons who are bound by long-term contracts stand to gain or lose by changes in prices. Thus, the farmer who purchases his farm subject to a mortgage when the price of wheat is $2 per bushel finds it difficult to make his required payments if the price of wheat falls to 50 cents per bushel. Persons whose money incomes are relatively stable are also greatly affected by price changes. The white-collar worker, whose wage or salary does not fluctuate much from year to year, finds that every increase in prices lowers the purchasing power of his money income, and every decrease in prices increases it. On the other hand, those whose incomes vary with changes in prices, such as entrepreneurs and speculators, may profit greatly by an increase in prices. They are generally able to charge higher prices for their products or services, and costs do not increase so rapidly as prices. As a result, each dollar received more than compensates them for the reduction in purchasing power. However, these same groups often suffer greatly in periods during which prices are declining; for prices often decrease more rapidly than costs,

and thus the loss in total dollars more than offsets the increase in purchasing power.

General price level and individual prices. There are two basic causes for the disturbances and losses caused by a change in the general price level. One, which we have already noted, is the existence of long-term contracts, especially debt contracts. The second is the fact that some prices are rigid and resist change more than others. This means that, when the general price level rises or falls, these rigid prices do not change in proportion. The prices of many goods in our economy are determined in highly competitive markets and are very flexible; the prices of other goods, however, are controlled by monopolies or the government and respond slowly if at all to changes in demand; others, again, are more or less fixed by custom. Because of these differences, any change in the general price level is accompanied by irregular movements of individual prices. This causes the whole system of price relations to be disrupted, with resulting loss or gain to various groups. We have pointed out, for example, that a rise in prices is likely to injure the white-collar worker. He is injured, however, not just because prices rise. The real trouble is that his salary, which is the price of his labor and is inflexible, does not rise so much as prices in general.

Business activity and the rate of spending. Up to this point we have emphasized the relation between money and prices, so much so perhaps as to give readers the impression that monetary factors disturb our economy solely because they disturb prices. This is by no means true. To be sure, where there are free and competitive markets any substantial change in money expenditures will quickly have an effect on prices. However, if prices were so rigid that they did not respond at all to changes in the money demand for goods, instead of the economic effects of monetary disturbances being decreased they would be even greater.

To illustrate, let us suppose that the people of a country spend 100 billion dollars for goods and services in a certain year. In the following year, however, a wave of caution sweeps over them, and they are willing to spend only 90 billion dollars. If prices are flexible and drop 10 per cent in response to the decline in expenditures, people can still buy as much goods as before, and production and employment will be maintained. If, however, prices are rigid and do not decline, 90 billion dollars will buy only 90 per cent as much goods as 100 billion bought the preceding

year. As a result there will be a sharp drop in production and employment.

It is clear, therefore, that changes in the rate of spending (*MV* in the equation of exchange) affect business activity in two ways: (1) by directly increasing or decreasing the quantity of goods sold and (2) by bringing about price changes.

Some economists believe that the difficulties and losses that arise from changes in the quantity and velocity of money are entirely unnecessary. They are convinced that our monetary system could be so managed that it would be a stabilizing factor in our economy. Many of these advocates of monetary management propose stabilization of the price level as an immediate goal. They believe that if we could not entirely prevent fluctuations in the price level we could at least greatly reduce them.

Summary

This chapter explains why changes in the quantity of money affect the price level. By means of the equation of exchange it was shown that, if the velocity of money and the volume of trade remain constant, a change in the quantity of money must have an immediate and proportional effect on prices; that is, if the amount of money in circulation should double, prices would double. The discussion pointed out, however, that such a direct relationship does not actually exist because the other factors in the equation, namely, the velocity of money and the quantity of goods, also vary, usually in such a way as to offset partially the effect of changes in the quantity of money. Thus the chief effect of more money, especially in the short run, may be a slower rate of turnover. In the long run, however, changes in the quantity of money exert most of their effect on the price level.

This basic relationship is of the utmost significance to the businessman, for he stands to gain or lose if prices increase or decrease. It is important, therefore, not only that he understand the implications of the equation of exchange, but also that he understand the methods that have been devised to measure price changes. In the discussion we have explained how index numbers are used to measure changes in the price level and also in the purchasing power of money. The businessman can study the indexes prepared by the Bureau of Labor Statistics and can adjust his policies to these changes.

Ideally, money performs the passive functions of measuring value and facilitating the exchange of goods. However, changes in its quantity and velocity cause it to assume an active role in disrupting business relationships. In large part, monetary factors disturb business through their effect on the

price level; however, they also disturb it by their direct influence on the volume of trade, production, and employment. In either case they increase business risks and bring unwarranted losses or gains to many. A demand has arisen, therefore, for the establishment of monetary controls, one of the objects of which would be to stabilize the price level.

At this point it is of most importance to understand that the quantity of money, the velocity of money, and the volume of trade all affect the price level; and that the price level in turn has important effects on business relationships.

Study Questions

1. What is meant by the value of money?
2. What is meant by the expression "general price level"?
3. What is the relationship between the general price level and the value of money?
4. What are the four major factors involved in the relationship between money and prices?
5. State the equation of exchange and tell the meaning of each symbol.
6. Why is the equation of exchange often called a truism?
7. Does the equation of exchange enable us to tell how much a given change in the quantity of money will affect the price level? Why or why not?
8. What is the real value of the equation of exchange?
9. State the equation of exchange in six terms, and tell what each means.
10. What is the quantity theory of money? Is it the same thing as the equation of exchange?
11. With the aid of the equation of exchange, trace the probable effects of doubling the quantity of money.
12. Would these effects be the same in the long run as in the short run? Explain.
13. Do changes in the quantity of money cause changes in the price level, or vice versa? Explain.
14. What are price index numbers?
15. Why are price index numbers sometimes called price relatives?

16. What relationship exists between a price index number and the purchasing power of money?

17. How is a rising price level likely to affect each of the following?
   a. People who owe long-term debts, like a mortgage on a farm.
   b. People whose money incomes are relatively fixed.
   c. Businessmen whose incomes arise from the sale of goods.
   d. The general level of employment and production.

EXERCISES

1. Suppose that the total amount of money in circulation is $1,000,000 and that its average yearly velocity is 20. If the total trade in "goods units" is 5,000,000, what is the average price of a goods unit?

2. Suppose that the number of goods units increases to 10,000,000 and all other factors remain the same. What is the average price of a goods unit?

3. Suppose that in a certain year M = $25,000,000, V = 4, and T = 5,000,000 goods units. The following year M = $50,000,000, V = 3, and T = 6,000,000 goods units.
   a. What was the original price of a goods unit?
   b. What is the new price?
   c. Why did the price of a goods unit not double when the quantity of money was doubled?
   d. Using the first year as the base, what is the price index for the second year?

4. The following data may be used for the construction of index numbers based on weighted price aggregates.

Commodity	*Units sold,* 1932	*Unit price,* 1932	*Unit price,* 1942
Wheat per bu	45,000,000	$1.00	$1.45
Cotton per lb	80,000,000	0.15	0.20
Butter per lb	10,000,000	0.40	0.50
Coal per ton	40,000,000	2.00	3.00

   a. What is the index number in 1942 based on 1932 prices?
   b. What is the index number in 1932 based on 1942 prices?

c. With 1932 as the base, what is the index of the purchasing power of money in 1942?

5. A laborer who received 50 cents per hour in 1932 was paid 60 cents per hour in 1942.

a. Assuming the same number of hours' work, was he better or worse off in 1942 than in 1932?

b. What wage per hour in 1942 would have given him the same purchasing power he had in 1932?

6. The following price index numbers are based on 1926. Translate each into an index number of the value of money.

Year	*Index Number*
1926	100
1929	95.3
1932	64.8
1939	77.1
1944	104

7. When index numbers are constructed by the aggregate method, they may be translated in terms of a new base. We simply divide the indexes of all other years by the index of the year that is to be the new base. Thus, if the index number in 1936 is 115, based on 1932, and we wish to shift the base to 1934 when the index is 110, to find the new index for 1937 we divide 115 by 110.

Suppose that we have the following index numbers based on 1932:

Year	*Index Number*
1932	100
1934	110
1936	115
1938	130
1940	135
1942	140

a. Translate the index numbers by placing them all on 1936 as the base.

b. Repeat, using 1942 as the base.

c. Why should we ever want to change the base on which we calculate index numbers?

8. A report stated that the purchasing power of the dollar gradually decreased between 1939 and 1946. Considering the purchasing power as

100 cents in 1939, it was 85 cents in 1942, 79 cents in 1944, and 65 cents in 1946.

a. Was the gold value of the dollar affected by these changes?

b. What effect did this change have upon landlords whose rents were frozen in this period?

c. How did this change affect workers' wages?

References for Further Reading

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 10.

Chandler, Lester V.: "An Introduction to Monetary Theory," Harper & Brothers, New York, 1940, Chaps. 1 and 2.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 24.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 22.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chaps. 19 and 20.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 14.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 21.

Spahr, W. E., and Others: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. I, Chap. 21.

Taussic, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 29.

*

8. Income, Expenditure, and Employment

In the preceding chapter the equation of exchange was used as the basis for analyzing the effects of monetary factors on our economy. This is the most basic approach to the problem. In recent years, however, the study of monetary forces in terms of the equation of exchange has been relegated to the background, and attention has more and more become centered on the size of the national income, on the immediate factors that cause it to rise or fall, on changes in the rate of expenditure, and on the ways in which such changes affect prices, production, and employment. This method of studying the flow of money and its effects on our economy is sometimes called the income and expenditure approach.

This newer approach to the study of the monetary factors in our economy has at least three definite advantages over the traditional quantity theory analysis based on the equation of exchange. First, it seems to come somewhat closer to economic realities because it employs such familiar terms as "income," "expenditure," "saving," "investment," "production," and "employment." Second, it is more directly related to the monetary aspects of the business cycle. Third, and for our present purposes most important, it is a great aid in enabling the student of economics to obtain a better picture of how the economy as a whole operates. In other words, it helps him to see the forest and not just the trees.

It should be emphasized, however, that the study of monetary forces in terms of expenditure and income is not a substitute for the traditional type of quantity theory analysis in terms of the equation of exchange. Rather, the newer approach supplements the older one. It carries somewhat further the analysis of the monetary factors in our economy and,

as already noted, it comes closer to explaining these factors in terms of the actual processes of economic life.

Relation of Income to Expenditure

A very obvious fact, but one frequently overlooked, is that the source of all money income is money expenditure. Whenever one person receives money, someone else must pay out an equal amount. As a result, for the economy as a whole, gross money receipts (MV in the equation of exchange) must always equal gross money expenditures (which can also be expressed by MV). Of course, it is not true that gross money receipts always represent net income to those who directly receive them. If a retailer sells a typewriter for $100, only a small part of this sum represents net income or profit to him. Most of it must be passed on to others to cover various types of costs, including the wholesale price of the machine. Eventually, however, the whole $100 becomes net income to someone in such forms as wages, salaries, interest, rents, royalties, or profits.

Although the ultimate source of all net income is expenditures by the final purchasers of goods and services, it should be emphasized that not all the goods purchased are for consumption. A very substantial proportion of them are for *investment.* Investment, as the term is used in this chapter, means spending money for the acquisition or production of new goods which are not consumed but which are held for future use, or as a source of future income. Expenditure for investment is just as effective in creating income as expenditure for consumption.

Since all income arises from expenditures, it follows that, for the total national income to remain stable at any given level, total expenditures for investment and consumption must remain stable at the same level. The conditions for stability of the national income may be represented as follows: expenditure for consumption + expenditure for investment = a constant total expenditure = the national income.[1] If total expendi-

[1] Strictly speaking the right-hand member of this equation should include, in addition to the national income, allowances for capital consumption including depreciation. At present, however, these amount to not much more than 5 per cent of the national income, and to simplify the exposition we shall disregard them. Note that the concept of national income most useful for the purposes of this chapter differs from that employed by the Department of Commerce and more nearly resembles what the Department calls gross national product. As we use the term, the national

ture falls, income tends to fall by the same amount. If total expenditure rises, income likewise tends to rise by the same amount. It follows that any change in consumption or investment, if not offset in whole or in part by a change in the other, will cause income to change in the same direction and by an equal amount.

Consumption, Saving, and Investment

In a normal year by far the larger part of the national income is spent for consumption, that is, for things like food, clothing, and personal services which are used up during the year in question.[2] However, a substantial part of the national income is always represented by savings and, if these savings are not spent, expenditure will fall below income and income will shrink. For expenditure to remain constant and equal to income, so that the level of income will remain unchanged, investment must equal saving. If investment falls below saving, expenditure and income will contract; if investment rises above saving, expenditure and income will expand.[3]

Investment will fall below saving with a resulting decline in income

income consists of three major elements: (1) the "savings" or undistributed income of corporations; (2) the disposable income of individuals; and (3) the tax income of government units. Taxes are included in the national income because they are spent for goods and services which presumably benefit everyone; however, that part of tax receipts used to make transfer payments (like veterans' pensions) should be excluded. Such payments are included in individual income and ought not to be counted twice.

[2] Although the greater part of the national income is normally spent for consumption, the proportion varies from year to year. For example, of the national income as calculated by the Department of Commerce, about 85 per cent was spent for consumption goods in 1940 as compared with only 77 per cent in 1946. During the war the percentage was even lower.

[3] In his "General Theory" Keynes says that investment and saving must always be equal. This results, however, from the special way in which he defines them. With the terms employed in their usual sense of money saved or money invested, there is no necessity that saving and investment should always be precisely the same. It might possibly be argued that as soon as an investment is made, even if the source of the funds is newly created bank deposit credits, an equal amount of income accrues immediately to people somewhere in the economy; and since this income cannot yet be spent, it necessarily becomes savings. However, because of slowness of accounting procedures and of arrangements for the transfer of goods, such income will not be *realized* until some time later. If by income we mean realized income,

if enough people begin to hoard, that is to say, if they decide to keep part of their incomes in the form of money in order to build up their cash balances.[4] If month after month and year after year the public should attempt to hoard part of the net national savings, the national income would steadily fall and the final result would be national impoverishment. Something not far removed from this actually happened when the great depression was developing in the early 1930's. When such a decline in national income gets under way, it will not stop until one of two things occurs: Either private individuals and business firms stop contracting their expenditures, or else the government steps in and offsets the contraction in private spending by increasing its own spending. Usually such an increase in government spending is applied to relief or public works.

For action by the government to have much chance of success the additional spending must come from one of two sources: the sale of government bonds to people with idle cash balances, or the sale of government bonds to the banks in exchange for newly created deposit credits. If the government obtained the additional funds by increasing tax rates, it would defeat its purpose, for the higher taxes would cause private spending to drop still lower.

If the government failed to take action to increase the total national expenditure, the shrinkage of incomes would continue as long as individuals and business firms kept on trying to hoard part of their savings in order to build up their cash balances. Eventually, however, the national income would reach such a low level that for the economy as a whole net savings would no longer be possible.[5] At this point, however, the public would be spending the entire national income for consumption,

there will usually be a lag between the making of an investment by one person and the receipt of the income that it creates for others.

[4] It has frequently been pointed out that if *everybody* should begin to hoard money the attempt would defeat itself. Everyone would cut down his expenditures, but money receipts would tend to fall equally fast, provided no new money were created by the banks or the government. It should be noted, however, that individuals with relatively large incomes might have some success in increasing their cash balances at the expense of the lower income groups, who would have to reduce theirs because of unemployment and the necessity of falling back on savings.

[5] Some people might still be able to save, but this would be offset by the fact that others would be obliged to meet part of their current expenses out of capital or by borrowing.

and since expenditure would then be maintained at a level equal to income, contraction of income could go no further.

Just as income will keep contracting as long as investment falls below saving, so it will keep rising when investment exceeds saving. For considerable periods of time it is quite possible for investment to exceed saving. There are two ways in which the community as a whole can invest more money than it saves out of currently realized income: (1) The extra funds may come from hoards (including bank deposits) held by corporations and wealthy individuals. (2) They may be created by the banking system in the form of new deposit credits lent to businessmen.

Propensity to Consume and Rate of Saving

Statistical studies show that the proportions of their incomes which people spend for consumption, and the proportions which they save depend in part on the size of their incomes. To this relationship between the size of income and the percentage spent on consumption, Lord Keynes applied the term *propensity to consume.* As real national income increases, the propensity to consume declines. In other words, people save a larger proportion of their income. Thus, savings increase at a greater rate than income. On the other hand, as real income falls, the propensity to consume increases, so that savings decline at a faster rate than income.

This relationship between the size of the national income and the percentage saved is one of the factors that set limits to cyclical fluctuations of business. In a period of recovery the national income rises because total expenditures rise. Expenditures rise because money invested exceeds money saved, the excess of investment over saving coming from hoards and from newly created bank deposits. The fact that expenditure and income are rising in itself creates optimism and encourages a high rate of investment. But as the national income rises, savings rise even faster, and if investment is to keep ahead of saving, it too must rise faster than income. If it fails to do this, saving will overtake it. Once that happens, expenditures stop rising, income stops rising, and as a result investment becomes less attractive. Very soon investment begins to fall below saving, expenditure declines, and the national income also declines. It is clear, then, that because of changes in the propensity to consume, business can continue to expand only as long as investment rises at a greater rate than income.

Just as changes in the propensity to consume may check cyclical expansion, so they may also check cyclical contraction. National income declines in a period of business recession because total expenditures for consumption and investment decline. But as income falls, saving falls even faster, until eventually it drops to the level of investment, at which point the contraction of total expenditures and income stops. It is, of course, conceivable that the decline in investment might keep ahead of the decline in saving until, for the economy as a whole, net investment disappeared altogether. In that case, as was noted earlier, expenditure for consumption alone would eventually equal the entire national income and contraction of income would cease.

Variations in Prices, Production, and Employment

So far in this chapter we have been concerned almost entirely with the relations between total expenditures [6] and total national income expressed in terms of dollars. However, it is quite clear that changes in total expenditures do much more than merely change the national income in dollars. They also cause substantial changes in prices and in the volume of production and employment.

A decrease in the money spent by the public for consumption and investment means a decrease in the demand for goods. As was pointed out in the preceding chapter, if money expenditures dropped 10 per cent and prices also dropped 10 per cent, the change would not be very important because the smaller amount of money would buy the same amount of goods and there would thus be no reduction in production or employment. There would of course be a 10 per cent decline in the national income expressed in money, but real income would remain unchanged.

Unfortunately, an equal and uniform decline in prices is not what would really happen if money expenditures declined by 10 per cent.

[6] As has been noted, total expenditures include those made by individuals, business firms, and units of government. At present, if we include transfer payments, not far from 25 per cent of all payments in the United States are made by governmental units. Transfer payments are those not made in return for current goods and services. They include items like veterans' pensions, subsistence payments to veterans in school, and social security benefits. Though such payments are not made for goods and services, they create income for individuals.

Prices would fall, but in a very irregular fashion. Some would drop readily, but others would decline little or not at all. Wage rates would almost certainly resist the downward pressure, and this in itself would prevent the prices of some kinds of goods from falling very much. The irregularity of the decline in prices would in itself, by disrupting normal price relationships, tend to reduce production and employment. In addition, however, the average price level would lag behind the fall in money expenditures. If expenditures fell 10 per cent, the price level might fall 5 per cent. The reduction in spending would therefore mean an over-all decline in sales, production, and employment; and national income would decline not only in terms of money but also in terms of real purchasing power.

Multiplier Effect

One of the inherent characteristics of a highly specialized free-enterprise economy is that any initial expansion or contraction in the rate of spending tends ultimately to result in a considerably greater expansion or contraction of both expenditure and income. This is commonly called the *multiplier effect.*

Suppose, for example, that in a certain month expenditures for the construction of buildings decline by 100 million dollars. The incomes of workers in the building trades will be directly reduced by a decline in wage rates or by unemployment, as will also the incomes of workers in the industries that furnish building materials. The reduction in the incomes of these workers will force them to reduce expenditures for consumption, and this will reduce employment in other industries and cause the incomes of more workers to fall. Before the multiplier effect is exhausted, the total contraction of income and expenditure may exceed by several times the original 100 million dollar reduction in expenditures for new buildings.

The multiplier effect can cause expansion as well as contraction, especially if an increase in expenditure occurs at a time of depression and unemployment. This is the principle behind the theory of "pump priming," or the theory that the government, by a certain initial expenditure on public works, can stimulate a recovery of private business that will become self-sustaining. In theory, pump priming is doubtless sound,

but there are certain considerations that restrict its successful application. One difficulty is to make the initial expenditure by the government large enough to overcome cyclical contraction of business and to get the expansion cycle well started. Another serious difficulty is to make sure that spending by the government does not arouse such fears on the part of businessmen that private investment is curtailed enough to offset the increased spending by the government.

Effect of Speculation on Expenditure, Income, and Employment

By speculation, as we employ the term here, we mean the tendency of people to increase their expenditures when they believe that prices will rise, and to decrease them when they think prices will fall. When prices are rising, people increase their purchases of goods and property rights for one of two reasons: (1) they may hope to sell in the future at higher prices; (2) they may hope to avoid paying higher prices at a later time. As soon, however, as prices begin to fall and the belief spreads that the fall will continue, people tend to reduce their purchases sharply because (1) they are now overstocked with goods and (2) they expect to be able to buy later at lower prices.

The speculative factor contributes a serious element of instability in a free-enterprise economy. The expectation of higher prices is the very thing that makes prices rise and keeps them rising. The increase in expenditures brings a rise in employment and income, and also a rise in inventories. Eventually, however, the price rise must come to an end, and as soon as the reversal in trend is recognized, expenditures, employment, and income begin to contract. Since lower prices are now expected, demand becomes sluggish and the result is that prices continue to fall. Eventually, of course, inventories will be reduced; and when the price level is sufficiently low, resistance will develop to a further decline. Then a new rise will start and the whole process will be repeated.

It should be noted that, when prices are falling, people prefer money to goods, or to use Lord Keynes's expression, their *liquidity preference* is high. They strive to hoard money and to build up their cash balances. On the other hand, when prices are rising they prefer goods to money and try to reduce their cash balances to a minimum.

Level of Employment

It is obviously desirable that expenditure and income should be reasonably stable at a high enough level to provide full employment. In our modern free-enterprise economy the greatest single economic problem is recurring periods of large-scale unemployment. Such unemployment constitutes a vast waste of productive power, creates untold hardships for workers, brings fear and uncertainty, and endangers the whole fabric of our political and economic life.

Large-scale unemployment is a problem that grows out of the depression phase of the business cycle. With each return of prosperity it either disappears or becomes reduced to minor proportions. However, except during a war and the period following a war when business firms and consumers are replacing their stocks of goods, business operates at prosperity levels for very limited periods of time. As a result, a substantial amount of involuntary unemployment is more or less the usual situation. At the bottom of a major depression it reaches huge proportions. In 1932 and 1933, for example, the National Industrial Conference Board estimated unemployment at between 12 and 13 million workers, or about 25 per cent of our entire labor force.

Until recent years nearly all economists have more or less assumed that, if the business cycle could be reduced to minor proportions by stabilizing expenditure and income, the problem of unemployment would solve itself. The theory behind this assumption was that, once wide fluctuations of income and employment were overcome, the "automatic" forces of the market—the forces of free competition—would then raise employment to a high level and hold it there. This assumes, of course, that the methods employed to stabilize business activity would not themselves interfere with the effectiveness of competition.

When competition is reasonably effective in the markets for goods and services, there is some foundation for the belief that it tends to raise the level of production and employment and hence the level of real income. Probably it raises the level of money income too, but that is less important. If prices, including wage rates, are flexible enough to respond readily to changes in supply and demand, adjustments take place which tend to expand output and employment toward capacity levels.

To illustrate how competition tends to increase employment, let us

assume a situation in which the level of production is insufficient to provide full employment. If free competition prevails in the labor market, the competition of unemployed workers for jobs will cause wage rates to fall, and to keep falling as long as unemployment persists. This fall in wages will increase the demand for labor in various ways. For example, it will make it possible to employ labor for some purposes for which its use was previously unprofitable. It will also encourage the construction of factories, machinery, and other capital goods because it will reduce their cost. Further, it will increase employment in factories because it will make it possible for manufacturers to increase sales and output by reducing prices. The manufacturers, on their part, if they are selling in a highly competitive market, will be forced to make price reductions in order to hold their position in the trade. In these ways and in others, competition tends to cause price adjustments which result in the expansion of production, employment, and real income.

In recent years, however, the whole area in which prices are freely determined by competition has been steadily shrinking, and increasing numbers of prices have been brought under controls of one type or another. Some prices are controlled by industrial monopolies or by tacit agreements among large producers; others are controlled by the manufacturers of trade-marked products; and still others are controlled by government. For many years government price controls were applied chiefly to the utility industries, but in recent times there has been a great expansion of government price controls outside of the utility field. In part these controls have been introduced to meet the problems of depression, and in part to meet the problems of war. There is yet another agency of price control that it would be a mistake to overlook. One extremely important group of prices, namely, wage rates, is coming more and more under the control of organized labor.

The net effect of this expansion of price controls has been to make our entire price system much more inflexible, and to weaken greatly those automatic forces upon which economists have traditionally counted so heavily for keeping production and employment at high levels, or for bringing it back to those levels in case of a cyclical contraction.

Today few economists are optimistic enough to believe that the forces of competition, left to themselves, can be counted on to prevent extreme contraction of employment or to assure reasonably prompt recovery, once contraction occurs. In view of what happened in the 1930's it would be

difficult to have such faith in the power of the "automatic forces" of the market. Moreover, many economists who are disciples of Keynes question the basic theory that the competitive forces of the market always tend to expand production until a level of full employment is reached. Keynes argued that, in a society where incomes were relatively high, a stable equilibrium might be reached at such a low level that involuntary unemployment would still persist.[7]

Inherent Instability of Our Economy

The economists of the nineteenth century paid relatively little attention to business cycles. They believed that the forces acting from within the economic system always tended to expand production in the direction of full employment. They attributed cyclical contractions of economic activity not to forces normally operating in the economy but rather to such disturbing "external" influences as bad crops or war. In the twentieth century, however, the increasing severity of depressions brought a radical change in attitude. Since the First World War cyclical fluctuations of business have been the great center of economic interest.

Obviously a war, or even a series of bad crop years, may have a powerful effect upon the level of business activity. However, it is pretty generally recognized today that the successive oscillations typical of modern business are not primarily the result of special disturbing influences. Rather they grow out of certain characteristics inherent in the kind of economy we have. There are, it is probable, certain underlying forces in our economy which are always tending to expand production and employment toward capacity levels. Nevertheless, in a specialized

[7] Briefly, Keynes's theory runs somewhat as follows: As the income of an economy rises because of technological progress, consumption rises at a slower rate while savings rise at a more rapid rate. Finally, production methods become so efficient that if full employment could be maintained there would not be sufficient investment opportunities to absorb all the funds that would be saved. Therefore, full employment is never reached (unless the government intervenes) because consumption plus investment cannot rise high enough to absorb the amount of product that would be turned out at full employment. Keynes's theories are still highly controversial. A fuller development of them might show that, insofar as they are valid, they require the assumption of certain rigidities in the price system. Keynes was not so much interested in long-run economic trends as he was in the problem of stabilizing production and employment at relatively high levels.

economy where almost everyone depends for an income upon selling goods or services in the market, innumerable circumstances may cause a temporary decline in sales and income. In most cases these declines would not be serious except for one thing: the nature of our economy is such that once a contraction gets well started it tends to be cumulative. Because of this, what might otherwise be a minor recession sometimes develops to major proportions.

The cumulative nature of business fluctuations is a result of various characteristics more or less inherent in the economic system. Among these are the behavior of bank loans and bank deposits; the effects of the speculative motive on the demand for goods; and the multiplier effect of a given expansion or contraction of expenditure.

Basic Conditions of Full Employment

It is clear that the problem of maintaining a satisfactory level of employment is twofold. On the negative side it is the problem of preventing serious cyclical contractions. On the positive side it is the problem of bringing production to a sufficiently high level to provide jobs for all who want them.

Reduced to its simplest terms, the solution of the problem of preventing cyclical contractions of employment is to find some means of stabilizing national income and national expenditure. This statement should not be taken to imply that the solution is easy, for in a free-enterprise economy it is extremely difficult to find practical ways of stabilizing either income or expenditure. We should not forget, however, that an effective way of *influencing* the flow of money (MV) in the form of both income and expenditure is to regulate the volume of money (M); and regulating the volume of money means principally controlling the volume of bank deposits.

In spite of the crucial importance of the rate of income and expenditure, we should not assume that control of the business cycle is a purely monetary problem. Income and expenditure are both monetary factors, but to a large degree they are controlled by other factors that are non-monetary in nature.

Given the price level and a reasonable degree of stability, from the positive side the problem of providing full employment is the problem of raising national income and national expenditure to a sufficiently high

level. If expenditure for consumption contracts, expenditure for investment must be increased. If necessary, methods must be devised for expanding investment or consumption, or both. Here again monetary measures are likely to be an aid, but they are not the whole answer to the problem.

If population increases and if we assume a given price level, income and expenditure must increase sufficiently to provide employment for the additional workers. If production increases, not because of a growth in population, but because of improved methods that increase the output per worker, one of two things must happen if employment is to be maintained. Either the price level must be allowed to drop gradually so that the existing volume of income and expenditure will maintain employment by taking more goods from the market; or else money expenditures and incomes must be raised to absorb additional goods at the already existing price level.

Even if fluctuations in business were largely ironed out, raising employment to a permanently satisfactory level might prove very difficult. It is, however, a basic problem that must be dealt with.

Summary

This chapter has carried somewhat further the analysis of the effects of monetary factors on our economy. The interdependence of national expenditure and national income was noted, as well as the effects that changes in expenditure and income have on production and employment. It was found that, as long as expenditure remains constant, income, production, and employment will remain constant; that is, provided there is no change in the price level or in methods of production. It was also found that national expenditure will remain constant so long as there is no tendency for investment to fall short of or exceed saving. If investment keeps falling short of saving, expenditure, income, production, and employment will fall. If investment keeps running ahead of saving, they will rise. Because of changes in the propensity to consume, people tend to save a larger portion of their incomes as incomes rise, and a smaller proportion as incomes fall. To some extent this fact tends to limit cyclical fluctuations in income and employment.

It was noted that a rise or fall of national expenditure would be a matter of little consequence if all prices changed in the same proportion. The disturbing effects of changes in the rate of expenditure occur principally because (1) the prices of individual goods are affected differently and (2) changes in

the general price level are less than, or lag behind, changes in the rate of expenditure. As a result, changes in the rate of expenditure have substantial effects on production and employment.

Two important factors were explained which tend to cause changes in the rate of expenditure and income to become cumulative. These were the multiplier effect and the speculative motive.

The desirability of full employment and the seriousness of the problem of cyclical unemployment were discussed. The tendency of the automatic forces of the market to expand employment to maximum levels was explained, but the conclusion reached was that these forces are weaker than formerly and that they obviously cannot be counted on to prevent major cyclical contractions of employment, or to bring reasonably prompt recovery. It was also noted that there are powerful forces inherent in the economic system itself which make for instability and sometimes bring about major contractions of expenditure, income, production, and employment.

Finally, it was pointed out that, expressed in the simplest terms, the problem of maintaining full employment is twofold: (1) cyclical fluctuations of income and expenditure must be reduced and (2) national income and national expenditure must be raised to adequate levels.

We have now completed our discussion of the effects of monetary factors on the economy. In the following chapter we shall turn our attention to an entirely different type of problem, namely, the determination of the prices of individual commodities. In order better to concentrate our attention on this latter problem we shall assume that monetary stability has been achieved and that the general level of incomes, expenditures, employment, and prices remains constant.

Study Questions

1. Show how the income and expenditure approach to the study of monetary factors can be related to the equation of exchange.

2. What are the advantages of the income and expenditure approach over the older quantity theory approach?

3. Explain the relation between national expenditure and national income.

4. If everyone should start to hoard money, what would be the effect on the total amount of money held by all individuals and business units put together? What would be the effect on national expenditures? On national income? On employment?

5. Could total money expenditures rise above total money incomes? Why or why not?

6. "Since part of the national income is saved, expenditures must always be less than income." Comment.

7. If the government decides to increase its own spending to offset a contraction of private spending in time of depression, should it raise the necessary money through taxes or by selling bonds to the banks? Explain the reason for your answer.

8. If investment persistently fell below saving, would the national income finally reach zero? Why or why not?

9. What conditions must exist with respect to expenditures for consumption, expenditures for investment, and total expenditures if the national income is to remain stable?

10. It is said that the propensity to consume depends on the size of the national income. Explain.

11. Explain how changes in the propensity to consume might check (*a*) a cyclical rise in business activity and (*b*) a cyclical decline.

12. Will a contraction of expenditure reduce prices or reduce production and employment? Explain.

13. What would be the economic effect of a decline in the national income of 90 per cent if all prices, including contract payments of interest and rent, fell in proportion?

14. What is meant by the multiplier effect? Explain why it operates.

15. Explain why the speculative motive causes income and employment to go through alternating periods of rise and decline.

16. What is meant by liquidity preference? How is it related to speculation?

17. What is meant by the automatic forces of the market? Explain how some of these forces tend to raise the level of production and employment.

18. Why are prices less responsive to competitive influences than they formerly were?

19. Why do economists have much less faith than formerly in the power of the automatic forces of the market to maintain production and employment at reasonably high levels?

20. Give some reasons for the inherent instability of the economic system.

21. What are the two somewhat different problems that must be solved if employment is to be maintained at a high level?

6. A theater is presenting a famous play for one night only. It can seat 1,800 people. The charge is $5 per seat, but, because of the popularity of the play and the cast, 3,000 people apply for tickets.

a. If the demand for seats has an elasticity of unity, how much could the management have charged and still sold all 1,800 seats?

b. Suppose that it had been planned to give the play two evenings. On the basis of the demand indicated above and assuming that no one wished to attend more than one performance, how much could the management have charged if it had wished to fill all the seats for both evenings?

7. At a world series game scalpers sold reserved seats for $25. The box office price was $5. Why did not the management set the box-office price at $25?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 12.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 8.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 10.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 9 and 10.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 7.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 17.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 14.

Marshall, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book V, Chaps. 2, 3, 5.

Spahr, W. E., and Others: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. 1, Chap. 16.

11. Cost of Production

The analysis in the preceding chapter has indicated that price is determined by the interaction of supply and demand. From the side of demand, the desire of individuals and their ability to pay for a good determine the number of units that will be taken from the market at any given price. From the side of supply, the willingness of producers to continue to bring goods to the market determines the amount of goods that will be offered for sale at any price. Thus the forces of supply and demand determine not only the equilibrium price, but also the quantities that will be purchased and sold.

Cost of production and price. As a rule, the most significant factor affecting the willingness of producers to bring goods to the market is cost of production. Under competition, if the average cost of a commodity is less than its selling price, the resulting profits will tend to induce an increase in production, and this in turn will tend to bring price down to cost. If the average cost of a commodity is greater than its selling price, the effect will be the opposite. The losses that firms suffer will bring about a decrease in production, and this will tend to bring the price up to the cost. There is, therefore, a strong tendency for the price of a commodity, under competition, to equal its cost of production. This general principle is, however, subject to important qualifications, which will be noted and explained later.

Meaning of cost of production. As used in economic analysis, the term "cost of production" includes all sacrifices that are made to produce a good or to render a service, insofar as these can be given a monetary value. Sometimes the monetary valuation will be arbitrary and may represent only a reasonable guess. Wherever possible, however, monetary valuation of sacrifices made should be determined on the basis of payments made by others for comparable sacrifices.

Let us suppose, for example, that we are to determine the monetary value of the sacrifices being made by a grocer who owns the building in

which his store is located, and who with the help of his wife operates the store. Obviously, the grocer knows the cost of the commodities that he buys and sells; but he does not know the amount of rent to charge for his building nor the amount of wages to allocate to his own efforts or those of his wife. To estimate rent and wages his most logical procedure is to discover how much rent is being paid in his community for a store similar to his own; and also the amount of wages others are paying for services similar to those which he and his wife are giving. Sometimes he can ascertain these monetary values by definite offers received from others who wish to rent his store or to secure his services. In any case, if he will make an effort to determine what his rent and wages would be if he were receiving them from others, he will be in a position to determine the cost of using his own land and his own services. This information will aid him in determining whether he can afford to stay in business. The tragedy is that many small businessmen never determine the money value of the sacrifices they make, and in consequence they have no way of knowing whether or not they are recovering all their costs of production.

The example given above illustrates the nature of the cost problem. It shows that all persons who give services to a business should be assigned a money wage equal to the market value of their services. Likewise, all other factors should be assigned a share even though they are owned by the businessman himself. If all such costs are carefully calculated, then those who produce goods or render services will be in a position to know the real cost of producing a given item. Business, however, is complex; hence, even if the most carefully constructed cost-determining system were used, every item of cost could not be found exactly. Later some of the difficulties involved will be considered. At present we are concerned primarily with defining the concept of cost as used in economics.

Cost of production includes all payments made to others for goods or services, as well as such items as depreciation and obsolescence. It also includes an estimated wage for services rendered by owners, plus a return to capital and land supplied by owners. Among the more important cost items are wages, interest on capital, rent, payment for raw materials, maintenance, taxes, wages of management, and depreciation. Against depreciation certain reserves are commonly set aside, and other reserves may be created for contingencies of various sorts. Such reserves, together

with the return on the labor, capital, and land supplied by the owners, do not represent definite payments to others; hence, they are not always recognized as costs. Costs, in fact, are often thought of as including only actual expenditures, and after such expenditures have been subtracted from total income, the remainder is called profit. Such "profits," however, include both reserves and the normal amount of interest that owners should receive on their investment. Therefore, from the economic point of view, profit calculated in this way includes items that are really costs.

This manner of estimating profits is confusing, since true profits cannot arise until all costs have been taken into account. Surely owners are entitled to a definite return for their services and for the capital that they have put into the business. This return should be comparable to that which they could receive if they invested in other ventures. Also adequate reserves should be set aside for depreciation; and something for such contingencies as obsolescence, although the latter is very difficult to estimate. True economic profits cannot arise until after these factors have been counted as part of the cost of doing business.[1]

It has been emphasized that economic costs include not only payments made to others, but also a fair return on the factors provided by the businessman himself. To the economist, therefore, no profits have been made until after all such costs are covered. Thus, if the economist says that a business has made a profit of a million dollars, he means that this amount is left over *after* wages, rent, and interest have been allowed to the owners. However, the usual method of calculating profit does not allow for all these factors. Corporations, for example, never count interest on the investment of their stockholders as a cost. As a result, a business often reports profits when in reality there are none.

An example that shows the confusion in regard to the term "profits" is found in a booklet issued by the Chrysler Corporation and containing remarks made by K. T. Keller, president, at a recent annual meeting of

[1] One serious problem arises, however, when allowances are to be made for capital invested in the business. A corporation, for example, may be started by promoters who issue securities which have a certain par value. However, the amount actually invested may not correspond with the par value of the stock. Promoters often allocate to themselves a certain amount of stock as payment for their services. Often, however, the par value of the stock is greater than the real value of such services. Certain other methods of obtaining stock may also cause the real investment to be less than the par value of the stock.

the stockholders. Under the heading, "Big Figures Don't Mean Fat Profits," Mr. Keller states:

> You run into big figures when you deal with big-scale manufacturing. The digits string out far behind the dollar signs. Sometimes people confuse these big figures with unusual profits—especially in wartime. . . .
>
> Now a billion and one-half dollars is a large figure to understand. For instance, it is more than the expenditure of the entire national government of the United States for the two years 1914 and 1915 combined. It is almost half again more than the value of the entire 1943 wheat crop of the country.
>
> This billion and a half is the money we receive for our war products. Out of it we pay all the people that work in our plants; out of it we have to pay for materials, manufacturing facilities, and supplies of all kinds; pay freight bills, pay social security and other taxes, take care of the cost of upkeep and other repairs, and pay all the everyday expenses entailed in turning out a large quantity of war items.
>
> Out of the money that comes in we also have to set aside some of it to get our machinery back in place and fix up our plants in order to make automobiles again. . . .
>
> It is only after we have taken care of such items as these—which use up most of the money coming in—that we have a little left to put back in the business or pay to the stockholders. So when you say Chrysler Corporation had a profit for the two war years of millions of dollars, what does it mean? Few people, if any, thought that Chrysler made big profits before Pearl Harbor. We were successful, yes. . . .
>
> Now let us compare profits before and after Pearl Harbor. In the five years prior to Pearl Harbor (1937-1941 inclusive) profits averaged $8.47 per share of stock per year—and this stock is the property of some 50,000 different people who own the Chrysler business.
>
> As against $8.47 average earning per share per year in the five prewar years, in the two years since Pearl Harbor (1942-1943) profits averaged $4.46 per share per year—a little more than half the profits in the prewar years.

Mr. Keller's statement illustrates well the fact that when businessmen think of "costs" they seldom include in the term all the economic costs of production. It is clear that if the stockholders invested any capital at all in return for their shares, neither the prewar earnings of $8.47 per share nor the war earnings of $4.46 were all profits in the economic sense. If we suppose, for example, that the stockholders invested $100 per share,

and that the normal rate of interest was 5 per cent, earnings would have had to be $5 per share to cover the interest cost on the investment of the stockholders. On this assumption prewar profits were only $3.47 per share, while during the war stockholders actually suffered an economic loss of 54 cents per share.

To determine whether in fact the earnings per share of stock were unreasonably large or small is not easy. First, we would have to know the amount of the real investment of the stockholders; then we would have to know the market rate of interest; and finally we would have to know whether all costs other than interest had been correctly estimated. If we knew these things, we could calculate the true economic profit per share. Even then, however, we would not have answered the question of whether the profits, if any, were "reasonable" or not. Men go into business not just to earn a fair rate of interest on their capital, but in the hope of making economic profits to compensate them for their risks. In our economy the hope of profits is a necessary incentive to induce businessmen to undertake new enterprises and expand old ones; and there is no simple rule to tell us just how large profits need to be to provide businessmen with an adequate incentive.

This explanation is, of course, no criticism of the Chrysler Corporation. Mr. Keller's statement is used simply because it illustrates the problems that are confronted when the question of profits is raised.

When the income that accrues to a business is carefully allocated so that fair returns to owners are included as cost, one can clearly see that the amount remaining represents a true profit and gives an accurate gauge of the success of the business. In the long run the production of a commodity can be maintained at a given level only if all the factors of production receive sufficient compensation to prevent them from shifting to other industries. *Cost of production,* therefore, includes a normal rate of return to all the factors that are necessary to create a good and bring it to the ultimate consumer.

Opportunity costs. The cost of any factor used in producing a commodity may be regarded as an *opportunity cost,* because the cost of the factor is determined in a market in which many other "opportunities" are bidding for its use. Other things being equal, if a firm manufacturing shoes is to obtain funds to build a new factory, it must offer at least as high a rate of interest as that offered by the producers of furniture, clocks, or automobiles. Likewise, if it wishes to secure labor, it must bid

against all the other industries that employ the labor that could be used in the making of shoes. The cost of a factor of production is not determined in any one industry. Rather, it is determined by the *total* demand for it in relation to the supply. Consequently, all producers tend to pay like prices for the opportunity of using similar factors. Opportunity cost represents the greatest amount any factor could receive from some other producer, and it is this amount that the businessman using it must pay.

Cost, prices, and incomes. Though it is common knowledge that costs of production must ultimately be paid out of the price received for the product, this fact is often forgotten. For example, all too frequently it is assumed that businessmen can pay any wage rate they desire, thus increasing or decreasing their labor cost at will. Actually, of course, they cannot in the long run pay a wage rate higher than that which can be met out of money received for goods and services sold. Neither, on the other hand, can they pay wages much less than those paid by other employers; for if they do, either they will not get workers, or those they get will be of very inferior quality. The only exceptions to this rule are to be found in periods of large-scale unemployment or in cases where an employer is more or less isolated so that he does not have to compete against other employers for his labor supply.

Even when people understand that costs can be met only out of the prices received for the product, they often fail to keep in mind that costs themselves are nothing but prices. Costs of production are simply the prices that the businessman pays for the things he must have in order to produce commodities or services. In a competitive market they are determined like other prices, by supply and demand. Though they are costs to the businessman, sooner or later they are income to someone else.

It is very important to understand that payments that are costs to one group become income to some other group. It explains, for one thing, why a *general* contraction in business expenditures must bring a corresponding contraction in national income. Likewise, it explains why a general expansion of business expenditures must cause national income to rise. Of course, it should also be understood that business expenditures are tied to consumer expenditures, since business firms can meet their costs only out of money received from the sale of commodities and services.

Short-run relationship of cost to price. The short run, it will be recalled, is a period too short to permit variations in the rate of production; or,

in the case of some farm products, it is the period during which supply is limited to the stocks already on hand, because these cannot be increased until the next crop is planted and harvested. Since costs of production affect prices principally by causing changes in the rate of production, it is clear that in the short run their influence on prices must be limited. *Past* costs are irrevocably spent and should have no effect at all on the actions of either buyers or sellers, or on the price. Each seller, for example, should seek the best price he can get. If it is relatively high, he will make a profit. If it is low, he will at least, by selling, reduce his loss. However, estimates of *future* costs should have some influence on prices even in the short run, because these estimates will affect the willingness of dealers to offer their stocks for sale. If future costs and therefore future prices are expected to be higher than those prevailing at the moment, dealers will tend to withhold their stocks and this will raise current prices; on the other hand, if future costs and future prices are expected to be lower, dealers are likely to offer their stocks for sale, even at a loss, and this will bring prices down.

It should always be remembered that under perfect competition cost can affect prices only by controlling supply. When supply is wholly independent of cost, cost can have no influence on price. There are at least two situations in which this is the case. First, there is the short-run market condition which is sometimes called *fixed supply*. Where stocks of a perishable commodity are in the hands of a large number of competing dealers, each must take whatever price he can get regardless of cost; for if he withholds his stock he will suffer a complete loss. The case of the farmers selling strawberries, mentioned in the last chapter, is a good illustration of fixed supply. A second situation in which cost has no influence on price is where goods are not reproducible. In this case supply is permanently limited by existing stocks, with the result that demand is again the active factor in determining price. Examples of non-reproducible goods are paintings by old masters and first editions of famous books.

Relation of costs to prices in the intermediate period. The intermediate period, it will be recalled, is a period long enough to vary production within the limits of existing plant capacity, but not long enough to build new factories and new machines, or to retire old ones from use without loss. In the intermediate period cost of production exerts a powerful effect on price, but there is not a complete adjustment of price to cost.

The equilibrium price of a commodity for this period may be either above or below its average cost to most firms.

The remainder of this chapter and much of the following chapter will be devoted to an analysis of the relation of costs to prices, production, and profits in the intermediate period. We shall be especially concerned with the pattern followed by the costs of the individual enterprise as it varies its rate of output. For the purposes of this analysis we must begin by defining and explaining certain cost concepts.

Fixed and variable costs. In carrying on his business a producer finds that, for the intermediate period, some of his costs are fixed while others are variable. *Fixed costs are those which remain stable in total amount regardless of changes in the volume of production.* Suppose, for example, a businessman borrows $1,000,000 and with this money purchases a building in which to conduct his business. Let us assume that the interest rate is 5 per cent. Under these assumptions the businessman has assumed an obligation to pay $50,000 per year in interest whether the building is being used or not. Such a fixed cost bears no relationship to the rate or amount of production. *Variable costs,* on the other hand, *are those costs which vary in total amount with every increase or decrease in output.* A good example of a variable cost item is raw materials. Suppose a businessman is producing footwear. Each time he produces a pair of shoes a certain amount of leather is required. Therefore the cost of the leather used will vary with production. That is why this item is called a "variable" rather than a "fixed" cost.

Certain costs that are fixed if output does not change too much may vary as production increases or decreases beyond certain limits. Thus if production increases enough, total wages of management may eventually increase; likewise depreciation, maintenance, and even taxes. However, these items do not vary in direct proportion to output; and in many instances no sharp line of distinction can be made in classifying fixed and variable costs. Nevertheless, the distinction between them is important, for while many costs vary directly in total amount with every change in the volume of production, others remain approximately constant in total amount over considerable periods of time and through a wide range of output.

Marginal cost. *Marginal cost* is another concept which is of special significance in the intermediate period. *The marginal cost of a product is the addition to total cost which results from increasing output by one*

unit. Marginal cost is of great importance to the businessman because it is the cost that he must consider when he weighs the advantages of expanding his output. The question he must answer is, Will the *additional cost* be less than or greater than the additional gross income? As long as the additional cost of producing another unit of product is less than the additional income, it pays the businessman to expand his output. This means that, under perfect competition, the businessman will make the greatest possible profit by expanding production up to the point where the marginal cost of his product is equal to the price that he receives. Even where competition is not perfect, this principle is a good guide if competition is so effective that the businessman has very little control over the price of his product.

Fixed and variable costs in relation to output, average cost, and profits. As production in a given plant is expanded from a very low to a very high level, we find that average unit costs at first decline and later begin to rise again. These changes in cost as the rate of production is increased can be explained chiefly in terms of two factors: (1) the relation of fixed to variable costs and (2) the principle of diminishing returns. We shall defer discussion of diminishing returns to the next chapter. At present we are interested in explaining how the relation between fixed and variable costs affects output, average cost, and profits.[2]

We have already noted that, as production in a plant is expanded from a low level, average unit cost at first drops. Let us see how this can be explained in terms of the relation between fixed and variable costs. Suppose that Mr. Smith, after surveying the market, estimates that if he goes into the shoe manufacturing business he can, within a reasonable period of time, sell 500,000 pairs of shoes a month at a profit. At first, however, his production and sales would have to be much less than this. In order to prepare for the demand that he foresees in the future, he builds a factory capable of producing 500,000 pairs of shoes under normal operating conditions. When he begins operations, he produces only 100,000 pairs of shoes a month, but from time to time he increases the rate of production; and as he steps up production his average unit cost drops. Let us see why this is so.

[2] The existence of fixed costs accounts in large part for the declining phase of the average cost curve of the individual firm. The rising phase of the cost curve is accounted for by the principle of diminishing returns. The relation of diminishing returns to the average cost curve will be explained fully in the next chapter.

To simplify the problem we make certain assumptions. First, we shall assume that as he increases production his *unit* variable cost for labor, material, and selling expenses remains constant. Second, we shall assume that his fixed costs are really fixed; that is to say, their *total* remains constant so long as he does not increase his present plant capacity. Third, we shall assume that these fixed costs, which include interest, taxes, maintenance, depreciation, and wages of management, amount to $100,000 per month. Finally, we shall assume that his raw materials, wages for labor, and other variable expenses total $2 per pair of shoes. We are now in a position to determine Mr. Smith's average unit costs of production on the basis of the size of his output. Let us see what these costs would be for the various levels of output as shown in Table 9.

Table 9

Amount produced	*Total fixed cost*	*Average fixed cost* (2) ÷ (1)	*Average variable cost*	*Average total cost* (3) + (4)
(1)	(2)	(3)	(4)	(5)
100,000	100,000	$1.00	$2.00	$3.00
200,000	100,000	0.50	2.00	2.50
250,000	100,000	0.40	2.00	2.40
300,000	100,000	0.33⅓	2.00	2.33⅓
350,000	100,000	0.28	2.00	2.28
400,000	100,000	0.25	2.00	2.25
500,000	100,000	0.20	2.00	2.20

The table shows that as Mr. Smith expands production toward the rated capacity of his factory, his average fixed costs decrease. This fact means that the average total cost of shoes will decrease since variable costs remain constant per unit of output. It is clear that, if he must limit his production to 100,000 pairs of shoes, he will take a loss if he has to sell them for less than $3 per pair. However, if consumers will take 500,000 pairs, he can sell the shoes for $2.20 per pair and still secure a return equal to his cost of production. We see, therefore, that the quantity of product that he estimates will be demanded is a very impor-

tant factor in determining the willingness of a businessman to undertake a new enterprise.

When a business firm expands production from a low level up to the capacity of its plants, as long as variable unit costs remain constant, average unit costs must decrease. The reason for this is the decline in average unit fixed costs. In the following chapter, however, it will be pointed out that variable costs per unit of product do not always remain constant. If production is pushed far enough, they will sooner or later begin to rise, and eventually they will rise very sharply.

The rapidity with which average unit cost will fall as production in a plant is expanded to rated capacity will depend in large part on the relative importance of fixed and variable costs. If fixed costs are large when compared with total operating costs, then as production is increased the rate of decline in unit cost is rapid; if fixed costs are comparatively small, the rate of decline will be slow. In an extreme case there might be no fixed costs at all. Suppose, for example, that Mr. Smith was able to rent a building and machines under a contract providing that his payments should be proportionate to his output; suppose, further, that all those connected with the management of the business should work on a commission basis determined by output. If all costs were variable, unit cost of production might remain almost constant through a wide range of output.

Some businesses are operated with very little need for fixed expenses or overhead, and unit costs do not vary to any appreciable degree with variations in output. In other enterprises, however, the great proportion of costs are fixed. In a hydroelectric plant, for instance, practically all the costs are fixed, since they represent investment in the dam and electric generating equipment. Obviously, because practically all costs are fixed, unit costs will decrease rapidly as output increases to rated capacity. Under such conditions, the producer will be willing to expand production even though prices must be reduced as output is increased. The only thing producers will try to insist on is that consumption be increased sufficiently as prices decrease to allow them to increase, or at least to maintain, their profit margin. In situations of this kind, decreases in cost of production are a vital factor in inducing producers to bring additional goods to the market.

Summary

As a rule, *cost of production* is the most important of the factors that determine the supply of a commodity, so much so that under competition there is a strong tendency for cost to equal price. As used in economics, cost of production includes a return at the market rate to all the factors of production; even those that pertain to the owners of a business. This means that for a businessman to make profits in the economic sense he must have something left over after he has allowed interest on his own investment, rent on his own land, and a salary for his own services.

It is important to remember that costs of production are really prices. They are the prices that the businessman must pay for the things he needs in order to produce commodities and services. Like other prices they are determined by supply and demand. To the businessman who employs factors of production, their prices are costs. However, to the people who sell these factors, their prices constitute income. Costs of production are sometimes called *opportunity costs,* because to obtain a factor of production an enterprise must pay at least as much as the other "opportunities" that are bidding for its use. Under conditions of perfect competition all users of a factor would pay the same price for it.

In the short run costs of production have only a limited influence on price. In the special case of *fixed supply* they have no effect at all. In general, the longer the period of time under consideration, the greater the influence of cost on prices. However, if a good is not reproducible, there is no relation whatever between cost and price, even in the long run. This is because cost of production affects prices only by controlling supply.

In the intermediate period, the relationship between costs and prices is somewhat complex. To understand this it is necessary to study the individual firm and to see how changes in the rate of production affect costs. However, in order to analyze the relation between costs and the rate of output, it is necessary to define certain basic cost concepts, especially *fixed costs, variable costs,* and *marginal costs.*

We find that as production is expanded in a given plant, if variable costs per unit remain constant, average cost per unit declines. This is because fixed costs are spread over more and more units. If fixed costs are a relatively large part of total costs, the decline in average cost is rapid.

In the following chapter on Proportionality, the principle of diminishing returns will be examined. This principle also plays a major part in explaining the relation between costs and the rate of output, and hence in explaining the relation between costs and prices.

Study Questions

1. As the term is used by economists, what does cost of production mean?
2. Businessmen and economists do not always mean the same thing by costs of production. Explain.
3. Define profits as the term is used in the text.
4. The "profits" referred to by Mr. Keller in the quoted passage may not be profits at all in the strict economic sense. Explain.
5. It is not always easy to determine the true economic cost of producing a commodity. Why?
6. What are opportunity costs?
7. What is the difference between variable costs and fixed costs?
8. As the rate of production in a given plant is increased, explain what will happen to average unit cost
   a. If fixed costs are very large and variable costs per unit of output remain constant;
   b. If fixed costs are almost negligible and variable costs per unit of output remain constant;
   c. If variable costs per unit of output keep rising.
9. Which of the following can be considered fixed and which variable costs?
   a. Interest on borrowed funds used to purchase machinery.
   b. Raw materials used in production.
   c. Wages of labor paid on an hourly basis.
   d. Wages paid on a yearly basis.
   e. Interest on capital invested in the business by owners.
   f. Salaries of the president and other executive and managerial officers.
   g. Maintenance.
   h. Depreciation.
   i. Commissions paid to salesmen.

Exercises

1. The price of the stock of the X Corporation is $50 per share. Out of net income for the year the company is able to declare dividends of $5 per share and to add to surplus a sum equal to $2 per share. Estimate the amount of profits earned per share in dollars and also as a percentage on the

investment. What reasons make it difficult to determine the actual economic profits of a corporation?

2. *a.* What are the annual costs of production in the following business? Mr. Smith, the owner and manager, has invested $1,000,000 in buildings and machinery. In order to devote his time to this enterprise he gave up a $25,000-a-year position with a large corporation. Taxes, depreciation, and maintenance amount to $100,000 per year, and the prevailing rate of interest is 5 per cent. Material and labor costs are $3,000,000.

b. Of the above items, which the economist would count as costs, which might not be so considered by the businessman? Explain this difference in viewpoint.

3. The capacity of a factory is 200,000 units. The investment in the plant is $1,000,000. Interest and depreciation together equal 10 per cent of the investment annually. Other fixed charges are $50,000. Variable costs are constant at 50 cents per unit of output.

a. Draw up a table showing total fixed costs, average fixed costs per unit, average variable costs per unit, and average total costs per unit when the factory is operating respectively at 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 per cent of capacity.

4. The following cost conditions prevail in a factory:

No. of units produced	*Total fixed costs*	*Unit variable costs*	*Total variable costs*	*Total costs*	*Average costs*
10	$100	$10			
20	100	18			
30	100	22			
40	100	25			
50	100	28			
60	100	30			
70	100	35			
80	100	38			

a. Complete the table.

b. On a graph plot and draw curves representing (1) average fixed costs per unit, (2) average variable costs per unit, and (3) average total costs per unit.

c. What are possible reasons for the increase in variable costs as production increases?

5. An advertisement stated as follows:

> Every schoolboy knows that the price of an article is the sum of the materials in it, wages needed to make it, taxes taken by the government, profit without which the manufacturer couldn't stay in business.
>
> Add them all together and you have the price of the product. Increase the size of any of them, and price goes up. Two plus two still equals four.

a. Do you agree with the definition of *price* as used in this statement? Rewrite this article using the correct term in place of price.

b. What assumption as to the relationship of price and cost of production is made in this statement? Is this assumption correct? Why or why not?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 12.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1942.

Cady, G. J.: "Entrepreneurial Cost and Price," Edwards Brothers, Inc., Ann Arbor, Mich., 1942, Chaps. 1 and 2.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 9.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 10.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 10 and 11.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 10.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 18.

Marshall, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book V, Chaps. 7 and 8.

NOURSE, E. G.: "Industrial Price Policies," Brookings Institution, Washington, D. C., 1938.

STIGLER, G. J.: "The Theory of Competitive Price," The Macmillan Company, New York, 1942.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chaps. 12 and 13.

12. Proportionality

We noted in the preceding chapter that for a given enterprise some costs of production are likely to be fixed in the intermediate period. This results from the circumstance that, for this period, some of the factors of production are fixed. New plant and equipment cannot be built; neither can old plant and equipment be abandoned without loss. As a result, expansion of production from a low level up to the rated capacity of existing plant causes average cost to drop by spreading fixed costs over more and more units of product. It should be emphasized, however, that fixed cost is essentially an intermediate-period concept, because in the long run all the factors of production used by an enterprise, and consequently all its costs, are variable.

In this chapter we are still interested primarily in analyzing the behavior of costs of production in the intermediate period. In the preceding chapter we assumed that as the production of a plant was expanded variable cost per unit of product remained constant. On the basis of this assumption, expansion of production caused average total cost to decline, though at a constantly decreasing rate, until the limits of plant capacity were reached. Actually, however, average variable cost is not likely to be constant. Up to a certain point it usually falls; beyond this point it rises. If production is pushed far enough, the rise in average variable cost will eventually more than offset the fall in average fixed cost, with the result that average total cost will also begin to rise.

Cost curve of the individual firm. We find, then, that the costs of an enterprise follow a characteristic pattern as production is expanded within the limits of existing plant capacity. If volume of output is plotted along a horizontal axis and average unit costs along a vertical axis, as in Fig. 1, a graph of average unit costs tends to be U-shaped. However, the precise shape of the average unit cost curve will vary greatly from firm to firm. In some instances the U may be nearly perfect; in others the base of the U may be flat, indicating a long range of output where

average unit costs are almost constant; in still other instances, the span of decreasing costs may be short, with increasing costs becoming evident very soon as output rises. Several possible shapes of this average unit cost curve are shown in Fig. 1.

We have noted that two factors account for the characteristic pattern followed by average costs. The declining phase of the U curve is explained

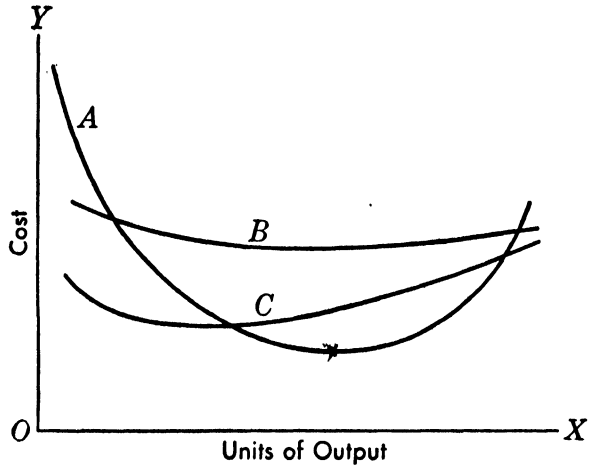


Fig. 1.—Variety of possible shapes of average unit cost curves as indicated by those for firms *A*, *B*, and *C*.

partly by the spreading of fixed costs over more units of product and partly by the tendency of unit variable costs to decline in the early stages of expansion. The rising phase of the U curve is accounted for entirely by the increase in unit variable costs which always occurs if production is pushed well beyond the most efficient capacity of a plant. Our next problem is to explain why variable costs per unit first fall and later rise as the rate of production is increased.

Principle of diminishing returns. The chief explanation of the early fall and later rise of unit variable costs is the *principle of diminishing returns.*[1] We have noted that, for a given firm, certain factors of production may be fixed in amount over considerable periods of time. In many cases the fixed factor is a manufacturing plant of a certain size. The firm cannot easily change the capacity of the plant; but it can readily

[1] Some writers prefer to call this law the *principle of diminishing productivity.*

increase or decrease its use of the variable factors, labor and materials, and so change its rate of production within fairly wide limits. However, because of the principle of diminishing returns, the amount of product obtained per unit of labor or materials will vary considerably with variations in the total quantity of labor and materials employed. The principle of diminishing returns may be stated as follows: *In a given period during which there is no change in the methods of production, if certain factors are held constant while the use of another factor is increased, up to a certain point the increase in output may be more than proportionate to the increase in the variable factor. However, after this point has been passed, the output will not increase in proportion to the increase in the variable factor.* Indeed, a point may even be reached where an increase in the variable factor will result in a net *decrease* in the amount of product.

Diminishing returns is a technological principle. It states the relationship between the physical quantities of the variable factor employed and the physical quantities of output. Its operation was first noted in agriculture, in connection with attempts to increase the output of a limited area of land by applying greater amounts of labor and capital; and its operation has been demonstrated time after time on experimental farms.

For the moment we are interested primarily in the theory of this principle, but many practical examples can be found of its operation in business and industry. For example, a businessman constructs a factory and equips it with a certain number of machine tools. That is the fixed factor. Next he hires men and purchases raw materials. Then men and raw materials are the variable factors. The raw materials flow into the factory, and the proprietor gradually increases his output. The law of diminishing returns tells him that his average output per unit of labor or materials will at first increase until a high point is reached, after which average unit output will start to decline.

The principle of diminishing returns is illustrated by Table 10, where we assume that a fixed amount of labor and equipment is to be employed on a specific piece of land. We also assume, however, that the amount of fertilizer used can be varied. The table shows the output that would be obtained in an average season if any one of various quantities of fertilizer were applied.

Let us analyze the results disclosed in the table. As successive units of fertilizer are applied to the land, the total product secured increases

constantly. However, this total reaches a maximum with the application of the eleventh unit, and the use of a twelfth unit of fertilizer actually decreases the total amount of product. We therefore call the application of the eleventh unit the *point of diminishing total returns.*

Table 10.—*Effect of Fertilizer on the Output of a Certain Piece of Land*

Units of fertilizer	*Total product*	*Total product added by fertilizer*	*Average product per unit of fertilizer*	*Marginal product of fertilizer*
(1)	(2)	(3)	(4)	(5)
0	10	0	0	0
1	13	3	3	3
2	24	14	7	11
3	40	30	10	16
4	74	64	16	34
5	100	90	18	26
6	124	114	19	24
7	136	126	18	12
8	146	136	17	10
9	154	144	16	8
10	160	150	15	6
11	164	154	14	4
12	154	144	12	−10

Column 3 shows the total amount of product that can be attributed to the use of fertilizer, and Col. 4 shows the average output added by a unit of fertilizer. Column 4 reveals that the average product of the fertilizer increases until after the sixth unit is used. The sixth unit produces the largest average product, and beyond this point average product declines. We call the point where average product is greatest the *point of diminishing average returns.*

Further information is revealed by Table 10. Column 5 refers to the marginal product. Before we go further, the meaning of the term "marginal product" should be carefully noted. It will be observed that,

when one unit of fertilizer is applied, the total added product is 3; whereas when two units are applied, the total added product is 14. Since the only condition changed is the amount of fertilizer, the entire increase in product must be due to the fertilizer. In other words, it is evident that two units of fertilizer make a better combination of the variable factor with the fixed factors than does one unit. As a result, the product is increased greatly. The term "marginal product" refers to the addition to total product that results from the addition of one unit of the variable factor. Thus, the addition of the second unit increases the total product from 3 to 14, an addition of 11. Again, the addition of the third unit increases the product from 14 to 30, or an increase of 16, and so on. It will be observed that the high point is reached when the fourth unit of fertilizer is applied, for the increase in product attributable to that unit is 34. Since this is the high point, it is called the *point of diminishing marginal returns.*

Two things must be noted: (1) The point of diminishing returns, whether total, average, or marginal, marks the end of increasing returns. (2) The points of diminishing returns are not all reached at the same time. In the illustration given above, the point of diminishing marginal returns is reached first; then average, and finally total. This is always the order in which these points are reached.

Diminishing returns and costs of production. Since in the illustration above, fertilizer is the only variable factor of production, it is clear that the price of fertilizer represents the variable costs of production. If the price of fertilizer never changed and if the average amount of product per unit of fertilizer never changed, average variable cost per unit would always be constant. For the present we shall assume that the price of fertilizer is fixed. However, we have already noted that, because of the law of diminishing returns, as production is expanded the average amount of product per unit of fertilizer first rises and then falls. Since the greater the product the less its cost, and vice versa, this clearly means that the average variable cost of the product first falls and then rises.

In the preceding chapter we illustrated by Table 9 how the existence of fixed costs affects average unit cost of production as output is expanded. To eliminate as much as possible the influence of variable costs, we assumed that these remained constant per unit of product. Now, however, our problem is to examine the behavior of variable costs and to discover how they affect average unit total cost.

To simplify our problem, let us first examine the behavior of variable costs alone. We used Table 10 to illustrate the principle of diminishing returns. We shall now use the same figures, with one modification, as a basis for illustrating changes in variable costs. The modification is this: Assume that no production at all is possible without fertilizer. As a result, the column headed Total product added by fertilizer will become the Total product column. Land, labor, and equipment are still the fixed factors, but to eliminate the influence of fixed costs assume for the moment that the fixed factors of production are free. All costs then become variable and result from the use of more or less fertilizer. Table 11 is based on these assumptions, and on the further assumption that fertilizer costs $20 per unit.

Table 11

Units of fertilizer	*Total product*	*Total cost*	*Average product*	*Average cost*	*Marginal product*	*Marginal cost*
1	3	$ 20	3	$6.67	3	$6.67
2	14	40	7	2.85	11	1.82
3	30	60	10	2.00	16	1.25
4	64	80	16	1.25	34	0.59
5	90	100	18	1.11	26	0.77
6	114	120	19	1.05	24	0.83
7	126	140	18	1.11	12	1.67
8	136	160	17	1.18	10	2.00
9	144	180	16	1.25	8	2.50
10	150	200	15	1.33	6	3.33
11	154	220	14	1.43	4	5.00
12	144	240	12	1.67	−10	

The table shows that the average cost at first decreases until a minimum point is reached, after which it increases, a result we would expect if we understand the principle of diminishing returns. The reader will note that, if we graphed the average costs shown in this table, we would get a U-shaped curve similar to the typical average unit cost curve which results when fixed costs as well as variable costs must be taken into account. This is to be expected, because the typical average total cost

curve, especially the rising portion of it, is to a considerable extent determined by the trend of average variable costs. Another fact may be noted in Table 11. If the marginal costs were graphed, the resulting curve would also slope downward at first and later rise; and it would intersect the average cost curve at the lowest point of the latter.

Principle of proportionality or best combination. We have now examined the effect of both fixed costs and the principle of diminishing returns on the average unit costs of a firm at various levels of output; and we have seen that these two forces account for the U shape of the average unit cost curve in the intermediate period. If certain factors of production are fixed and the prices of all factors are known, there will be a certain level of output, that is, a certain combination of the variable factors with the fixed factors, which will result in the lowest average cost. In practice, however, the average unit cost curve may be almost horizontal over rather wide variations in output. In any case, we must not fall into the error of supposing that the businessman should always stop expanding production when he reaches the point of lowest average unit cost. If often pays him to increase production even though his average cost is rising. This is true as long as the increase in cost is more than offset by increased receipts from the sale of the product.

We have seen that, where the production of a commodity can be increased by applying more of a variable factor to certain fixed factors, there is a point at which the physical output per unit of the variable factor will be greatest. If, however, we are thinking of the amount of product in relation not only to the variable factor but also to the fixed factors, there is no combination that is "best" in a purely physical sense. Even after production per unit of the variable factor has begun to decline, it will still increase in relation to the fixed factors; and if the fixed factors are especially scarce and valuable, it may be much more important to get a large per unit output in terms of them than to get a large output in terms of the variable factor.

It is only in an economic sense that there is clearly a "best" proportion in which to combine variable factors with fixed factors. In one sense the best combination is that which gives the lowest average unit cost for the product. To find this combination, however, it is not enough to know the quantity of product that the factors of production will yield when combined in various (physical) proportions. It is also necessary to know the prices that must be paid for the different factors.

In a sense even more significant to the businessman, the best combination of the factors of production is that which will produce the greatest profits. However, to find this we need still another kind of information, namely, the price that the producer can obtain for his product. We have already pointed out that it pays to increase production, even though average cost is rising, if the rising cost is more than offset by the price obtained for the additional product.

It is not difficult to cite examples where the entrepreneur would gain from producing beyond the rate of lowest average unit costs. During the Second World War, especially during its early stages, many war plants were operated beyond the level of lowest unit cost because volume of production, rather than cost or effort, was the most important consideration. Subsequently, when more war plants could be built, the output could be adjusted better to the level of greatest efficiency. It was no longer necessary to run the existing war plants at maximum output. In many cases the adjustment meant sharp reductions of average unit costs.

The entrepreneur is interested in maximizing his profits; hence, he wants to use that combination of factors which will bring him the highest net return. When the businessman can count on a fixed price for any volume of output, it often pays him to sacrifice high profits per unit for a greater volume at lower per unit profits. Even when he must lower his price in order to sell an increased output, he may find it desirable to expand production beyond the point of lowest unit cost, depending on the relation between costs, selling price, and volume of sales.

In the intermediate period it may pay a businessman to keep on producing even though he has to take a loss. It is conceivable that, even if an enterprise should expand production to the point of lowest average cost, average cost would still be higher than the price. Why, in this case, produce at all? The answer is that there are fixed costs which would go on anyway. Provided the product can be sold at a price that will more than cover variable costs, it is better to keep on producing because this makes it possible to cover part of the fixed costs. If production were stopped entirely, none of the fixed costs would be met, and the total loss would therefore be greater.

A study of Table 12 will give us a clearer understanding of the circumstances that a producer must take into account in deciding how far to expand production in the intermediate period. In this table it is

assumed that land is the fixed factor and that capital and labor are variable factors. This is somewhat more realistic than to assume that only a single item like fertilizer is variable. It is also assumed, in order to simplify the exposition, that when capital and labor are added, they are added in composite units containing a certain amount of each. To this last assumption there may be some theoretical objections, but it will serve our present purposes reasonably well. Finally, it is assumed that the piece of land used rents for $5 a year, that a unit of labor and capital costs $6, and that the product sells for $1 a unit.

Table 12 shows that the point of lowest average cost is reached when the sixth unit of labor and capital is applied. The least cost combination is, in this case, approximately 35 cents per unit produced. If only five units of labor and capital are applied, average cost is 39 cents per unit of product. If seven units of labor and capital are applied, average unit cost is 37 cents. Thus, average unit cost is greater if either more or less than six units of labor and capital are used.

The farmer now observes his position. His problem is: Shall I continue to increase production or shall I be satisfied when I have attained lowest average unit cost? His answer always lies in the relationship of marginal cost to the increase in receipts from sales. The increase in gross receipts that results from the sale of one more unit of product is known as *marginal revenue.* As long as the farmer's marginal cost, that is, the cost of producing one more unit of product, is less than his marginal revenue, it pays him to increase output. When marginal revenue and marginal costs are equal, he will be at the most advantageous position attainable. At that point his profits will be greatest, or his losses will be the least that he can expect.

In our particular example, as illustrated in Table 12, the problem is simplified by the fact that the farmer can sell any amount he produces at $1 per unit. Hence his marginal revenue is always $1. The farmer is selling in a market where competition is nearly perfect. In such a market marginal revenue is equal to price because the sale of additional units of product increases gross receipts by the full amount of the price received. We can say, therefore, that in a perfectly competitive market the entrepreneur will always find it advantageous to continue to increase his output until marginal cost is equal to price. It should be noted, however, that in order to find the point where marginal cost and price are exactly equal we would have to divide the factors of production and

Table 12.—Costs of Production of the Individual Firm as Output Expands

Factors of production		*Output*			*Costs*						*Income*	
Fixed	*Variable*	*Number of physical units of product*			*Fixed*	*Variable*	*Total*	*Average*	*Marginal*		*Total*	*Net profit or loss*
Land	*Number of applications of labor and capital*	*Total*	*Average* (3) ÷ (2)	*Marginal* (3)	*Land at $5 an acre*	*Labor and capital at $6 per unit*	(6) + (7)	*Per unit of product* (8) ÷ (3)	*Total* (8)	*Average per unit* (10) ÷ (5)	*At $1 per unit of product* (3) × $1	(12) − (8)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	1	3	3	3	$5	$ 6	$11	$3.67	$6	$2.00	$ 3	$−8
1	2	14	7	11	5	12	17	1.21	6	0.55	14	−3
1	3	30	10	16	5	18	23	0.77	6	0.37	30	7
1	4	64	16	34 ‡	5	24	29	0.45	6	0.17	64	35
1	5	90	18	26	5	30	35	0.39	6	0.23	90	55
1	6	114	19 †	24	5	36	41	0.35 §	6	0.25	114	73
1	7	126	18	12	5	42	47	0.37	6	0.50	126	79
1	8	136	17	10	5	48	53	0.39	6	0.60	136	83
1	9	144	16	8	5	54	59	0.41	6	0.75	144	85 ‖
1	10	147	14.7	3	5	60	65	0.44	6	2.00	147	82
1	11	148	13.5	1	5	66	71	0.48	6	6.00	148	77
1	12	144 *	12	−4	5	72	77	0.53	6	....	144	67

* Point of total diminishing returns.
† Point of average diminishing returns.
‡ Point of marginal diminishing returns.
§ Point of lowest average cost.
‖ Point of greatest profits.

the product into extremely small units and then increase production very gradually. In an arithmetic illustration like Table 12, where we add capital and labor in units of some size, it would be only an accident if, on the addition of a certain unit of the variable factors, marginal cost were exactly equal to price. Instead we are likely to find a point at which the addition of one more unit of capital and labor will cause marginal cost to jump from a level below the price to a level above it. In this case the producer should expand production only as far as is possible without pushing the marginal cost above the price.

In Table 12 we observe that average marginal cost remains below the selling price until the application of the tenth unit of labor and capital. However, with the application of this unit the marginal cost jumps to a level much above the price. Consequently, the producer represented by the table should employ only nine units of labor and capital.[2] It may be noted that when he employs nine units his total profit is greatest, and that the use of a tenth unit causes it to decline. It is clear, then, that employing nine units of labor and capital gives the best combination of the variable factors with the fixed factor.

We have already noted that in the intermediate period it may pay an enterprise to produce at a loss rather than to shut down completely, since, by producing, the loss may be reduced. In this case the principle that production should be increased as long as the *marginal* cost is less than the price still applies. If the *additional* cost of producing another unit of product is less than the price that can be obtained for that unit, it is clear that producing the unit will somewhat reduce the total loss that the enterprise is suffering. A firm that is losing money may reach the point where marginal cost rises above price *before it reaches the point of lowest average cost.* In that case it does not pay to increase production to the point of lowest average cost.

We can learn something more from inspection of Table 12. Marginal cost reaches a low point at the fourth application of labor and capital; nevertheless, since marginal cost is less than average unit cost, average cost continues to decline. However, once marginal cost becomes greater than average unit cost, average unit cost begins to rise. At precisely the point of lowest average cost, marginal cost and average cost are equal. The equality of marginal cost and average cost cannot be seen in the table because of the size of the units employed. If we could divide labor and

[2] This statement assumes, of course, that units of labor and capital cannot be divided.

capital, as well as the product, into extremely small units, we could find a point at which marginal cost and average cost would be exactly equal.

We have been assuming in discussing the relation between costs, volume of production, and price that the price is constant. Under perfect competition, price cannot be influenced by an individual producer; nevertheless it may go up or down as a result of changing conditions in the market. A study of Table 12 will make it clear that a change in price, if sufficient, will change the level of output that will yield maximum profits. A rise in price will induce the producer to increase production, even though such an increase will raise both marginal cost and average cost. Conversely, a fall in price will induce him to contract production.

Principle of variable proportions. Before we conclude this chapter on proportionality, one other principle should be stated, namely, *the principle of variable proportions.*

The principle of variable proportions is essentially a generalization of the principle of diminishing returns. The early writers on economics who stated the law of diminishing returns were thinking of land as the fixed factor; therefore, they assumed that the principle applied only to the product of labor and capital employed on land. Later it was recognized that under some circumstances *any* factors of production might be fixed. Consequently, the principle was stated in the more general terms of the definition given earlier in this chapter. This definition recognized that, no matter which factors are fixed, the principle of diminishing returns will apply to the factor or factors that are variable.

However, the principle of variable proportions goes even further than this. For most enterprises the assumption that some of the factors of production are fixed is valid only for the intermediate period. In the long run, for a given firm, all the factors of production are variable, and any of them can be either increased or decreased. Suppose, for example, that young Mr. Ford is considering his long-run policies in the building of automobiles. He can choose to use more machinery and fewer men, or he can choose to use less machinery and more men. Whichever combination he chooses, he can attain his goal, which is, let us say, the production of 1,000,000 automobiles per year. However, one of these combinations may be a considerably cheaper means of producing automobiles than the other, and obviously Mr. Ford will choose the one that he believes will produce that number of cars at the lowest average unit

cost. In doing so he is applying the principle of variable proportions. This principle may be stated as follows: *Various combinations of capital, land, and labor may be employed to achieve the same goal, but some of these combinations will result in lower costs than others.*

Just as there is no one best proportion in a physical sense in which to employ the factors of production when some are fixed and some are variable, so likewise there is no one best proportion, physically speaking, in which to combine them when they can all be varied. The best combination of the factors is that which will give the lowest cost, and in order to find this one must know not only the amount of product to be obtained by various combinations, but also the prices to be paid for the different factors. Since these prices vary in different countries, the best combination in one country may not be the best in another.

In the United States, for example, business has constantly sought to invent laborsaving machinery. Ours is the most industrialized nation in the world in terms of capital employed per unit of labor. In some other countries, like China, capital is very scarce, interest rates are high, and as a result few machines are employed. During the war, different methods of accomplishing the same objective often were used in China and the United States. When airfields were built in China, thousands of coolies were employed. By sheer sweat and toil the soil was removed and the field was built; the combination was a great quantity of labor and an insignificant amount of capital. Yet, because the cost of capital was very great and wages very small, the total cost of building the airfield may have been less than if bulldozers and other capital equipment had been employed, with only a few men to operate them. In the United States, on the other hand, when airfields were built, great quantities of capital equipment were used with relatively few men. In this country such a combination was best. It resulted in the lowest costs, because capital is cheaper here than in China, while wage rates are much higher.

We have pointed out that for a given enterprise, although some factors may be fixed over considerable periods of time, in the long run all factors are variable. For the economy as a whole, however, variability is much less. Even in the long run some factors of production are pretty definitely fixed in total quantity. For the most part these fixed factors are in the group that we classify as natural resources, or land. Any attempt, therefore, to increase the total product of the economy by employing more labor and capital brings into operation the principle of diminishing

returns. The effect of diminishing returns may be offset, and so obscured, by improvements in technology. If, however, methods of production remained much the same, the amount of product per unit of labor or capital would decline.

It is important to remember that the principle of variable proportions applies to those situations where *essentially the same results* can be obtained by using the factors of production in varying proportions. In certain industries the methods of production used determine rather definitely the proportion in which the factors must be combined. Over long periods of time some variation is possible, but only within rather narrow limits. In other industries, however, it is possible to vary over a wide range the proportions in which the factors of production are combined.

Summary

This chapter continues the analysis of the cost of the individual firm in the intermediate period. It notes that, as a firm increases production, its average cost curve tends to be U-shaped. The declining portion of this curve can be accounted for in part by the existence of fixed costs, the effect of which was explained in the preceding chapter. The chief purpose of the present chapter is to explain the changes in variable costs per unit of product which occur with changes in output, and to show how these changes in variable costs affect average unit costs, marginal costs, profits, and the rate of output itself.

If we assume that the prices of the factors of production are constant, changes in unit variable costs can be explained only by the principle of diminishing returns. This principle relates to the proportions in which the variable factors of production are combined with the fixed factors. It indicates that, as production is increased by employing more of the variable factors, in the early stages of expansion both the average and marginal output of the variable factors will increase. This will reduce marginal cost and tend to reduce average total unit cost. On the other hand, in the later stages of expansion both the average and the marginal output of the variable factors will decline, and this will raise marginal cost and tend to raise average total cost.

We conclude, then, that the declining portion of the typical average unit cost curve is accounted for partly by the existence of fixed costs and partly by the principle of diminishing returns, while the rising portion of the curve is accounted for entirely by the effect of diminishing returns on average variable costs.

The principle of diminishing returns deals with the relation of the physical input of the variable factors to the physical output of the product. Nevertheless, if we consider the fixed factors as well as the variable factors, there is no best combination in a purely physical sense. The best combination is the most profitable one. To find this we must know both the prices of the factors and the prices of the product. We should note, however, that the lowest average cost may not mean the greatest profit. Under perfect competition the greatest profit is obtained when production is increased until marginal cost is equal to price.

The principle of variable proportions is a generalization of the principle of diminishing returns. In the intermediate period, for the individual firm, some factors of production are fixed. In the long run all factors can be varied, and the producer must decide how much of each he will use. Often the same result can be obtained by employing the factors in different proportions; however, these different proportions will result in different costs of production. Again there is no best physical combination. The best combination is the one that gives the lowest average unit cost.

Though in the long run all factors of production are variable for a given firm, this is not true for the economy as a whole. Natural resources are more or less fixed in total quantity. Therefore, under any given methods of production, increasing output by employing more labor and capital will increase costs by reducing both the marginal and the average product of a unit of labor or capital.

Study Questions

1. Describe the usual pattern followed by average unit costs when production in a given plant is expanded from a very low level up to maximum capacity.

2. The principle of diminishing returns accounts in large part for the characteristic shape of the average unit cost curve of a firm. Explain why this is true.

3. How will the average unit cost curve of a firm be affected by the existence of large fixed costs?

4. Name the three points of diminishing returns. What is the order in which these points are reached?

5. Is there a best combination of the factors of production in a purely physical sense? Explain.

6. In an economic sense, when are the factors of production combined in the best proportions?

7. What kinds of information must the businessman have in order to find the best economic combination of the factors?

8. In a competitive market it often pays a producer to expand production even if in so doing he increases his average unit cost. Why?

9. When does it cease to be profitable to increase the rate of output of a factory?

10. Under what conditions will a producer find it in his interest to stop expanding production when he has reached the point of lowest average unit cost?

11. State the principle of variable proportions.

12. Why is the best proportion in which to combine the factors of production in one country not always the best in another?

13. There are limitations on the power of the businessman to change the proportions in which he employs the factors of production. What are some of these limitations?

Exercises

1. A producer has 1 acre of land on which he raises corn. The amount of corn he can produce depends on how many units of labor and capital he employs, as indicated by the following schedule:

Units of Labor and Capital	*Bushels of Corn*
1	10
2	25
3	45
4	62
5	70
6	75
7	78
8	77

a. Construct a table that will show, in terms of the number of units of labor and capital employed, (1) total product, (2) average product, and (3) marginal product.

b. On the table show: (1) the point of total diminishing returns, (2) the point of average diminishing returns, (3) the point of marginal diminishing returns.

2. Suppose that the acre of land in Exercise 1 rents for $4 per year, that a unit of labor and capital costs $5, and that a bushel of corn sells for 50 cents.

a. By means of a table show for the various levels of output: (1) average cost per bushel, (2) marginal cost per bushel, (3) total cost, (4) gross sales income, (5) total profits.

b. How many units of labor and capital should be employed to obtain the lowest average cost? How many to give the greatest total profit?

c. Compare average and marginal costs at the point of lowest average cost. Is this the point at which they are most nearly equal?

d. Represent the average and marginal cost curves on a graph.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 6.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 5.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 6.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 11.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 5.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 5.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 2.

Taussig, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 13.

13. Price in the Intermediate Period

Thus far our study of price determination has revealed that in the short run the amount of goods offered for sale can be varied only out of stocks on hand. The demand for such goods, therefore, largely determines the price at which they can be sold. Moreover, since the cost of producing the goods has already been undergone, such costs can have little influence in the market.

In this chapter we shall be concerned with another phase of price determination. We shall survey the forces that affect prices in the intermediate period; that is, the period within which plant capacity is fixed, but a period sufficiently long so that producers can increase output within the limits of plant capacity or decrease it if the price declines. The assumption of perfect competition will be continued. Certain initial changes in demand will be postulated. However, it will be assumed that during the period of price adjustment both the supply and demand schedules for a given commodity remain static or unchanged.

Demand schedules in the intermediate period. No sharp line can be drawn between demand schedules for the intermediate period and demand schedules for the short run or the long run. This is because the intermediate period is distinguished from the other periods by differences in conditions of supply, and not by differences in demand. However, demand schedules for longer periods of time tend to differ somewhat from those for shorter periods. It will be recalled that a demand schedule is a list of the quantities of a good that buyers are willing to take at various prices, in a given market in a given period of time. It is assumed, for the period in question, that the willingness and ability of buyers to purchase the commodity remain constant, except as affected by changes in the price of the commodity itself. In short periods, like a day or a week,

the action of buyers is likely to be greatly influenced by their opinions of future prices, and this fact is reflected in the short-run demand schedule. If buyers think the price of a commodity is going to rise next week, they will buy more this week and so temporarily build up their stocks. On the other hand, if they think the price will fall next week, they will buy less now in order to take advantage of the lower price in the future; and by this action they will temporarily deplete their stocks. However, buyers can build up or deplete their stocks only within limits. Over longer periods this factor becomes less important, and therefore expectations of future prices have less power to influence the demand schedule. As a result, in the intermediate period (and even more in the long run) the quantity of a good that people will purchase at any given price tends to correspond with the quantity that they are willing to consume at this price. It is assumed that their tastes and incomes, and the prices of other goods, remain constant. If any one of these factors changes, it will mean a change in the entire demand schedule.

Supply in the intermediate period. In the intermediate period the production of a commodity can change only as a result of a change in the rates at which individual producers operate their plants. If the price declines, some or all producers will reduce their rate of output and prices will rise. On the other hand, if prices increase enough as a result of an increase in demand, all the firms that are able to do so will step up output.

We must remember, as we view any intermediate period, that some producers may be operating near the physical limits of their plant capacity. Since, by assumption, no new plant capacity can be introduced, these producers are unable to increase output. Of course they may decrease it if prices decline. Other firms may be able to increase production only slightly. However, some producers with unused capacity can increase production greatly if prices climb. In the intermediate period it will nearly always be possible to step up the total output of an industry if a rise in prices creates a sufficient inducement.

If the price falls, each producer in the intermediate period can reduce output; and since a fall in price will bring price below marginal cost, each producer will be induced to contract output in order to bring down his marginal cost.

Let us observe more carefully how individual producers in an industry will react to a change in price. Suppose a firm with excess capacity finds

that its marginal cost is less than price. This firm will increase output to the level where price and marginal cost are again equal. Each producer will respond in the same way. As a result, additional goods will be produced and brought to the market. Contrariwise, each firm will reduce its output if it finds that marginal cost is above price. As a result, output of the industry will decrease.

We may conclude, therefore, that in the intermediate period total output will be increased or decreased as price rises and falls. Later in the chapter this relationship will be discussed further.

Demand for the product of the individual producer. In a market in which competition is perfect, the individual producer will receive the same price whether he produces much or little. In other words, from his point of view the demand schedule for the product will be perfectly elastic. The *market* demand schedule for the commodity will, of course, show that larger amounts can be sold only at lower prices. However, since any amount that the individual can produce will be negligible in comparison with total production, his contribution will not affect the price. Consequently, he will act as if he could sell an unlimited amount at the market price.

Figure 1*A* shows the market supply and the market demand for wheat in the intermediate period, and also the market price; Fig. 1*B* shows the demand curve and the marginal and average cost curves of an individual producer.

In Fig. 1*A*, *DD* is the demand curve and *SS* the intermediate period market supply curve. The equilibrium price is *PN*, or $1. In Fig. 1*B*, the price line is the demand curve of an individual producer, *AC* his average cost curve, and *MC* his marginal cost curve. Figure 1*B* shows that the individual producer receives a price of $1 per bushel no matter what his output. This price is determined by market conditions of supply and demand as indicated in Fig. 1*A*. It will be noted that when the producer in Fig. 1*B* expands his output to 10,000 bushels of wheat, his average cost and marginal cost are both 75 cents. At this point his average cost is at the lowest possible level. However, since every additional bushel of wheat can be sold for $1, it will pay this farmer to expand production further. If he follows the principles laid down in the previous chapter, he will continue to increase production until the cost of an additional bushel just equals the selling price. In other words, he will expand production until his marginal cost is just equal to $1. At this point his total

output, as indicated in Fig. 1*B*, will be 12,000 bushels; also, his total profit will be at a maximum. It is clear, then, that every producer of a commodity will attempt to expand production until his marginal cost is equal to the selling price. It follows that *in the intermediate period the marginal costs of the individual and in fact of all producers will tend to be equal to the price.* However, the reader should not assume from this

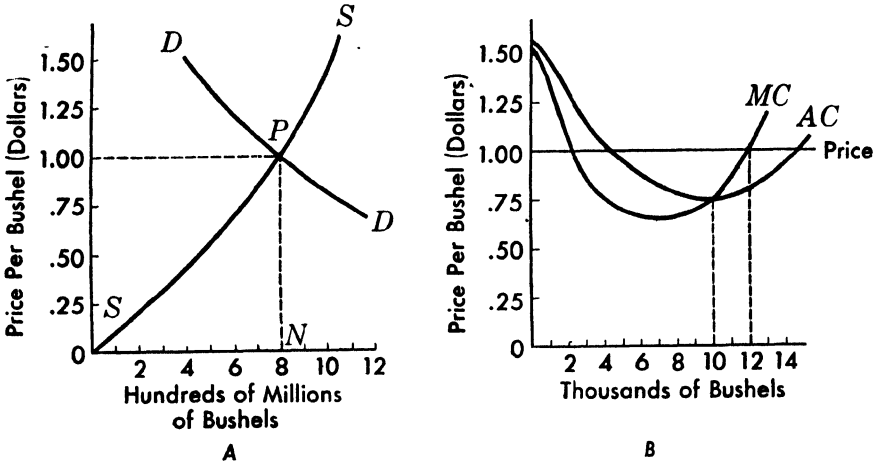


Fig. 1.—Relationship between market price and the demand and cost curves of the individual producer under perfect competition.

statement that marginal costs *determine* price. Cost of production is a very important factor in price determination, but it affects supply only. To understand how price is determined, the demand schedule must always be taken into account.

Equilibrium price in the intermediate period. The action of the individual producer gives us a clue as to the total supply that will come to the market in the intermediate period. Although, as has been pointed out, no one producer can by himself have any significant effect upon total supply, yet it is his output when combined with that of all other producers that affects prices.

Since any one producer increases production until his marginal costs equal price, if we wish to find the total quantity that will be offered at any given price, it is but necessary to total the outputs of all individual producers at this price. This output will be determined by a study of the marginal cost curve of each producer. Figures 1*A* and 1*B* show this rela-

tionship. In Fig. 1*A* the intermediate-period supply curve *SS* is derived from the marginal cost curves of all the individual producers.

Each point on the supply curve *SS* shows the total amount of wheat that would be raised and offered for sale annually if the price indicated prevailed. At high prices greater quantities will be produced than at low prices. This is because each producer will expand output until his marginal cost is equal to the price, and the higher the price the greater the expansion required to reach this point of equality. For example, if the price were 50 cents, when all producers had increased output until their marginal costs were also 50 cents, the total output of wheat would be only 500 million bushels per year. On the other hand, if the price were $1.50, when all producers had expanded until their marginal costs were equal to this amount, the total output of wheat would be 1 billion bushels.

However, the actual equilibrium price depends on demand as well as supply. The demand curve *DD* in Fig. 1*A* shows the quantities of wheat that buyers would purchase and consume in a year at various prices. The chart indicates that *PN*, or $1, is the intermediate-period equilibrium price. At that price the quantity that producers will raise and offer for sale is 800 million bushels; and this is also just the quantity that buyers will purchase and consume.

Differences in average costs. It should be noted that in the intermediate period the *average* costs of different producers may show variations even when price is in equilibrium. As a result, some producers may be making profits, some just breaking even, and others taking losses. Even so, the same principle applies, namely, that to make the largest profit (or the smallest loss) every producer should expand output until marginal cost equals price.

Effect of price changes on the individual firm. A change in the selling price of a commodity will, of course, change the profit situation of each firm. At a relatively high price a firm may be able to make substantial profits; at a somewhat lower price it may just break even; at a still lower price it may suffer substantial loss. Moreover, if the price is below average variable costs, the firm should not produce at all, since the variable costs can be saved by shutting down completely. However, if the price will more than cover variable costs, it is better to go on producing, even though the price will not cover average unit total costs. Any excess over variable costs helps to meet fixed costs; if the plant were shut down, fixed costs would go on anyway and would be a complete loss.

Figure 2 shows what the position of an individual firm might be at different prices for its product. *ATUC* is the average total unit cost curve, *AVC* the average variable cost curve, and *MC* the marginal cost curve. *PH*, P_1G, P_2F, and P_3E represent four possible prices of the product. The figure shows that, if the price were *PH*, this firm should expand production to *OH* in order to maximize profits. At this output marginal cost would equal price, but average total unit cost would be considerably

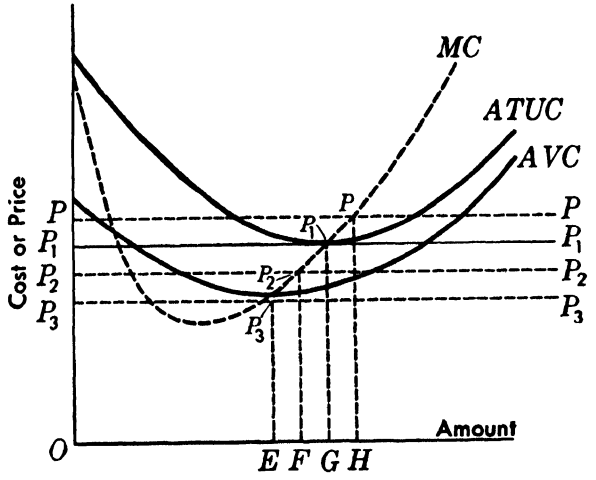


Fig. 2.—Position of individual firm when its product sells at any one of several prices.

below the price; and as a result profits would be substantial. If the price were P_1G, the firm should carry production to *OG*. At this output marginal cost, average total unit cost, and price would all be equal, and the firm would just break even. At a price of P_2F production should be *OF*. At this price it is not possible to cover average total unit costs, and consequently the firm must produce at a loss. However, it is better to produce than to shut down, for the price P_2F more than covers variable costs. Finally, at the price P_3E it would not pay to produce at all, since this price is less than average variable costs even at *OE*, the most favorable level of output.

Nature of the tendency of marginal cost to equal price. It is important to understand just what is meant when it is said that marginal cost tends to equal price in the intermediate period. The effective operation of this principle depends on two assumptions: (1) that producers are seeking

maximum profits and (2) that they are able to make reasonably correct judgments as to the prices that they will be able to obtain and the costs that they will incur. However, in a changing world in which the future is usually unpredictable, producers cannot possibly know just what *future* costs and *future* prices will be. Even under perfect competition, for producers to know future costs and future prices, the economic conditions affecting these costs would have to be *static;* or else the rate at which each condition was changing would have to be regular and predictable. In a static economy, or in an economy in which future costs and prices could be estimated with accuracy, there would be an exact equality between marginal costs and prices.

Summary

This chapter has explained the determination of prices in the intermediate period, under conditions of pure competition. Emphasis was placed on the conditions that affect supply. Further attention was given to the position of the individual firm in order to determine the conditions under which it would expand or contract production. This contributed to an understanding of supply because, although no single firm can influence the market, conditions that affect one firm may affect many; and expansion or contraction by a large number of firms will change the market supply and hence the price. It was found that *in the intermediate period each firm tends to increase production until its marginal cost equals price. Consequently, the marginal cost of all firms tends to equal price in the intermediate period.*

However, marginal cost does not determine price. Costs of production are the principal factor controlling supply, but to understand price determination we must also take into account demand. In the intermediate period both prices and the level of marginal costs are determined by the equilibrium between supply and demand. It was noted that, even when price is in equilibrium, in the intermediate period the average costs of different producers may vary. It was also noted that the equality of marginal costs and price is only a tendency. For this equality to be completely achieved, not only would competition have to be perfect, but economic conditions would have to be so static, or so predictable, that producers could accurately forecast both their costs and their prices.

Study Questions

1. In what respects is the intermediate period different from conditions found in the short-run period?

2. What conditions are necessary to cause an economy to become static?
3. Contrast demand in the short-run and in the intermediate period.
4. Contrast supply in the short-run and the intermediate periods.
5. Can total output be increased sufficiently in the intermediate period to have an appreciable effect on price? Explain.
6. Can the increase in output caused by any individual producer have any appreciable effect upon prices? Why or why not?
7. Why must the position of the individual producer be considered when the intermediate period is studied?
8. State the law of price in the intermediate period.
9. How is the individual firm affected by price changes?
10. Why does marginal cost tend to equal price in the intermediate period?

Exercises

1. In the intermediate period, would you expect to find increasing, decreasing, or constant average costs as an industry increased production to capacity? Explain.

2. Chart the marginal cost curve of an industry as production is increased to present capacity. Explain this curve carefully.

3. Suppose that a certain producer discovers that in the intermediate period the market price is below his average costs of production when he is operating beyond the point of least average cost. What should he do? Why?

4. Suppose the price is below his average costs after he has made all proper adjustments necessary in Exercise 3. Under what conditions can he still continue production and at what point should he stop production?

5. If competition prevails and no great progress has been made in ways of doing things, do you believe that the following statement is true? "In the intermediate period the price may be in equilibrium even though some firms are making great profits." Explain carefully.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, pp. 265-271.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 22.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chap. 15.

MARSHALL, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book V, Chap. 5.

14. Price in the Long Run

Nature of the long-run period. The long run, as has already been noted, is a period sufficient for a firm to adjust itself to new conditions by increasing or decreasing plant capacity. In some industries, where little skill and only a few simple machines are needed, operating capacity can be increased or decreased rather quickly as demand changes. Many small businesses are operated with a minimum of fixed capital and can spring into existence or liquidate with every increase or decrease in demand. Here the long-run period can sometimes be measured in weeks or months.

On the other hand, some industries require a large amount of fixed capital and an elaborate organization with skilled labor. Often such industries must operate on a large scale in order to reach optimum size and output, and for them the long-run period may stretch into years or even decades. The coal industry, for example, has for many years been unable to adjust itself to changed demand conditions; so also the railroads. In such instances, therefore, the period required for adjustment is long; and hence, when mention is made of long-run conditions, a period of years is involved.

Prices in the long-run period. To understand price determination in the long run it is necessary to analyze long-run cost trends. The analysis of costs in the intermediate period indicated that in this period all firms tend to produce until their marginal costs are equal, since each produces until its marginal costs coincide with the price in the market. In the intermediate period, therefore, the price of a commodity tends to equal marginal cost. In the long run, however, marginal cost is less significant, because given sufficient time each producer can increase the size of his plant and thus increase output without raising his marginal cost. If the price of the product is above his average cost, he has an incentive to expand output by increasing his plant capacity; and if enough producers increase capacity and output the price of the product will fall. On the

other hand, if the price of the product is below the average cost of a producer, and remains below cost, eventually he will be forced out of business; and if enough producers are forced to withdraw from the industry, eventually the price of the product will rise. In the long run, therefore, price tends to equal average cost of production for the industry, in the sense that it cannot remain much above or below the average cost of the typical producer.

Tendency of average costs to become equal. Under perfect competition, not only will long-run price tend to equal average costs for the industry; there will also be some tendency for the average costs of different firms to become equal. There are several reasons for this. First, over a long period of time there is likely to be a more even allocation of the factors of production among different firms. In the intermediate period some firms may have lower costs than others because they have superior factors of production; for example, machines. But if perfect competition prevails in the markets for the factors, each firm has equal access to the factors and will pay the same price for factors of equal quality. As a result, in the long run any differences in the quality or kind of the factors of production will tend to disappear.

A second reason that in the long run the average costs of all firms in the industry tend to become equal is that each firm has time to expand its plant capacity to the best, or optimum, size. In a given industry this optimum size is likely to be about the same for all firms. This is because in the long run and under competition no factors of production are fixed, and each firm will tend to have equal access to the factors and will pay the same prices for factors of the same quality. The reader will recall that in the intermediate period the average cost curve of the individual firm is U-shaped. This results from the fact that in the intermediate period plant capacity and certain costs are fixed. The long-run cost curve of the individual firm may also be U-shaped, but for different reasons. The shape of the long-run cost curve depends largely on changes in efficiency that result from changing the size of the plant. In many industries an increase in plant capacity brings economies of large-scale production. There is, however, an optimum size; and if plant capacity is expanded beyond this optimum, costs will no longer fall but will begin to rise.

Figure 1 represents the relationship of several intermediate-period cost curves to the long-run average cost curve of a certain firm. *LC* represents long-run average cost as the firm expands output by increasing

the size of its plant. A_1C_1 is the intermediate-period average cost curve when the firm operates a plant of less than optimum size; A_2C_2 is the intermediate-period cost curve when it operates a plant of optimum size; and A_3C_3 is the intermediate-period cost curve when it operates a plant of greater than optimum size.

The long-run average cost curve, *LC*, represents the lowest cost at which any given output can be produced after the firm has had time to make all possible adjustments. Consequently, no part of any intermediate-

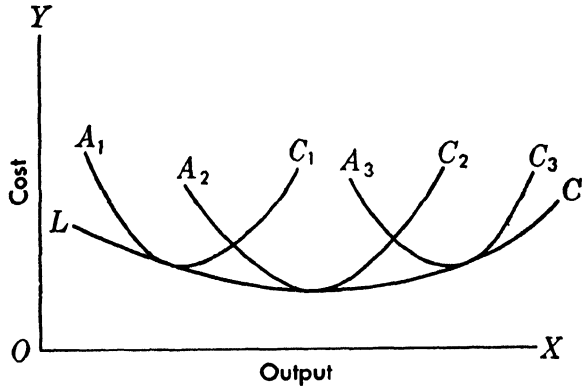


Fig. 1.—Average costs in the intermediate period and average costs in the long run.

period average cost curve can lie below *LC*. The long-run cost curve, *LC*, must envelop, or be tangent to, the intermediate-period curves. This means that when average long-run cost is declining, it is more economical to build a larger plant and utilize it partially than it is to build a somewhat smaller one and operate it at the point of lowest average cost.

Where competition is free, in the long run each firm in an industry will attempt to expand plant capacity and output to the point of lowest average cost. Each firm is obliged to do this in order to survive because the competition of other firms tends to bring the price of the product down to this lowest average cost. Hence the long-run price of a commodity tends just to cover economic costs of production when these are at the lowest possible level.

In the real world of business, competition is not perfect, and many factors interfere with equalization of the costs of different firms. However, if we remember that, when we refer to costs we mean economic

costs, we will find that some persistent differences in costs are more apparent than real. Often, for example, an enterprise will seem to have unusually low costs simply because the owners have failed to allocate the proper rate of return to the factors that they themselves own.

To illustrate, let us suppose that the wheat market is in long-run equilibrium with the price at $1 a bushel and average total costs are also at $1 a bushel. However, a certain farmer finds that his average costs are only 80 cents. Being intelligent, he determines to find out why he is apparently able to produce at costs below the general average. Perhaps he discovers that his soil is more fertile than most soils, or that it has unusually good drainage. If so, it is obvious that he is in possession of land that is more valuable than the land owned by his neighbors. His land is more valuable because of its productive power. A careful analysis will cause him to reappraise the productive contribution of his factors. He will then assign a larger share to land as a rental cost which should come to him as the owner; and in consequence his cost of production will no longer appear to be below the $1 a bushel which is that of his neighbors. Another possible explanation of costs that appear to be low is a better combination of the factors of production. In this case, since the enterpriser is responsible for combining the factors, he should allow himself higher wages of management as a reward for his superior ability. Here again what appear to be lower than average costs disappear when correct allocations are made to the factors of production.

It should be noted, however, that, even under pure competition and static conditions, the tendency for all firms to have the same average costs would never completely work itself out, unless we define "static conditions" so as to exclude any change in the circumstances affecting each individual firm. Usually, however, when we speak of static conditions in an economy we refer only to the general conditions of economic life. The constant factors in a static economy would be such things as population, the tastes and general mode of life of the people, the resources available, and the methods of production. Even in a static economy there would have to be life and activity and some types of change. Individuals would be born, grow old, and die, even though the general character of the population remained the same. Likewise new firms would be organized from time to time, and old firms would fail and go out of business. Although general conditions and average costs in an industry might remain constant, the circumstances affecting individual

firms would change in unpredictable ways, and their fortunes would rise or fall as their management passed from one group to another. As a result, even under perfect competition in a static economy, the average costs of all firms would never become exactly the same. The differences, however, would be much less than we actually find them in the world of business.

Long-run cost trends in different industries. When an entire industry changes its rate of production in the long-run period by changing its plant capacity, long-run average costs are also likely to change. However, the long-run average cost curve for an industry may be quite different from the intermediate-period average cost curve for the same industry. This is because the forces that operate in the long run are different from those which operate in the intermediate period.

In the intermediate period if an entire industry expands production from a very low level to a very high level of present capacity, it is likely to show a U-shaped average cost curve, similar to the average unit cost curve of the individual firm. This pattern is a familiar one. As output is expanded, average cost at first falls; later it may for a time be approximately constant; but eventually, if production is pushed far enough, average cost rises. These cost trends for an entire industry in the intermediate period simply reflect the cost trends of the individual firms that make up the industry. They result from the circumstance that in this period some factors of production are fixed for each firm in the industry, and therefore for the industry as a whole.

The long-run cost curve for an industry, however, may follow a different pattern from that just described, because in the long run there are no fixed factors or fixed costs for the individual firm. Given time, each firm can increase its plant capacity or expand its use of any of the factors. This means that it can expand output without bringing into operation the principle of diminishing returns.

To explain the long-run relation between costs and output as an entire industry expands production by expanding plant, it is necessary to assume that all the factors that affect costs, except plant capacity and the rate of output, are held constant. For example, even if there is no change in capacity and output, costs will fall if methods of production are improved. In order, therefore, to study the relation between cost and the long-run volume of output, it is necessary to assume that methods of production remain unchanged; or at least it is necessary to assume that there are

no changes in methods of production except those which result directly from changes in plant capacity and in the volume of output. The problem is to discover how such changes alone will tend to affect average costs of production in an industry.

When the long-run relationship between output and average unit cost is investigated, great differences are found between industries. Some industries, as production is expanded over the long-run period, are subject to *increasing cost;* others are subject to *constant cost;* and still others are subject to *decreasing cost.*

Increasing-cost industries. When expansion of production brings a long-run increase in the average costs of an industry, it is most commonly because of a limitation in the total supply of one or more of the factors of production. For an individual firm all factors of production can be varied in the long run. For an entire industry, however, this is likely to be true only within limits. If an industry is an important user of any factor, expansion of output by the industry will increase the demand for this factor and hence its price. Moreover, there are definite limits to the total supplies of certain factors, in particular those which we classify as natural resources, or land.

The best illustration of an industry of increasing costs is farming, especially if we think of farming as a single industry instead of breaking it up into such subdivisions as hog raising, corn raising, and cotton raising. However, when we say that farming is an industry of increasing costs, we do not mean that the actual costs of producing agricultural products must necessarily rise. In fact the very opposite may be the case, for improvements in agricultural methods may result in substantially lower costs. All we mean is that, if the general methods employed in farming remain unchanged, any expansion in the total output of farm products will, *in itself,* tend to raise average costs of production.

It is quite clear that farming is an industry of increasing costs. Let us suppose, for example, that a rise in the prices of agricultural products induces the farmers of the country to expand production substantially. If we assume that no general improvement in agricultural methods takes place, this expansion can occur in only two ways: (1) Farmers may cultivate more land; but for the most part this will be poorer land, which it did not pay to cultivate before, and costs of production on it will be higher. (2) Farmers may raise more on the land already in use by cultivating it more intensively; that is, they may use on it more of such

things as labor and fertilizer, factors of production that are less limited in supply than is land. However, when they intensify cultivation, eventually the law of diminishing returns will operate, and as a result costs of production will rise.

To illustrate the tendency toward increasing costs as land is cultivated more intensively, let us consider the situation faced by a farmer who decides to increase production on his land by using more fertilizer. The land has a certain natural fertility to begin with, and the farmer may be in a position to estimate what his crop will be if he applies no fertilizer. He knows, however, that an increase in fertilizer wll increase his output, and he can therefore experiment to determine the effect of an increase in fertilizer upon production and upon average unit costs. Table 13 shows the production per acre that could be obtained by adding fertilizer. We shall assume in the problem that, before any fertilizer is applied, the land will produce 60 bushels of corn per acre and that fertilizer is applied in units of 100 pounds. We shall also assume that fertilizer sells for 10 cents per pound and that other costs aside from the fertilizer are fixed and amount to $30 per acre.

Table 13

Units of fertilizer applied	*Total production*	*Total fixed cost*	*Total variable cost*	*Total cost*	*Average cost per bushel*
0	60	$30	$ 0	$30	$0.50
1	65	30	10	40	0.62
2	70	30	20	50	0.70
3	80	30	30	60	0.75
4	90	30	40	70	0.77

The table shows that, although the farmer is able to increase production with every application of fertilizer, his average cost per bushel increases as output is expanded.

Constant-cost industries. In the long run the costs of some industries tend to remain constant as output is increased. Constant-cost industries usually meet two conditions: (1) They can increase their use of the

factors of production without causing the prices of the factors to rise much. This is either because they depend principally on factors that are relatively plentiful, like labor, or because they use a very small part of the total supply of any factor. (2) An increase in their total output does not have much effect on the methods used in production. The hand-painted china industry would tend to have constant costs for both of these reasons. Again, though farming in general is clearly an industry of increasing cost, lettuce growing taken by itself may well be an industry of approximately constant cost. Doubling the output of lettuce might simply mean using twice as much land, labor, and capital without any change in the methods of production. There would, to be sure, be an increased demand for land, but it would be so small that it would have no appreciable effect on either land prices or land rents. Likewise there would be an increased demand for labor and capital, but again it would be too small to affect either wage rates or interest rates. As a result, the average costs of producing lettuce might be almost unaffected even though production were increased 100 per cent. It must be apparent, then, that many industries meet rather well the conditions that allow them to operate under constant cost.

Decreasing-cost industries. In some industries an increase in the total volume of production will, in the long run, result in lower average costs per unit. This is especially true, up to a certain point, of industries in which the use of machinery is important. In such industries, plants must be very large if full advantage is to be taken of the savings possible from the use of expensive, highly specialized machines. Until total production reaches the point where at least one plant can be operated at lowest unit costs, the costs for the industry as a whole will decline with expansion of output. This decline is a result of what are sometimes called *internal economies*, that is, economies within the organization of the individual plant.

When an industry is large enough to support plants of optimum size, no further internal economies will result from continued expansion of output. It is quite possible, however, that certain *external economies* will be realized by the industry and that these will bring a still greater reduction of average unit costs. External economies are those that result, not from improved organization of the individual plant, but from the growth of the industry as a whole. The cost of specialized machines, for example, is likely to go down if their manufacturers can sell them in

larger numbers. Accessory industries may be established to provide some of the things that the industry needs. Finally, as production increases, technical schools, trade journals, and various types of specialized services may appear, and as a result average costs may drop even lower.

It seems probable that many industries are industries of decreasing costs only in the earlier stages of their development. Once an industry is large enough to take full advantage of the gains from large-scale production and to be well supplied with service and accessory industries, further growth is less likely, in itself, to bring economies. There is little doubt that the automobile industry was once an industry of decreasing costs. It is not certain, however, that it still belongs in this group. The question is, Would a further increase in the capacity and average annual output of the automobile industry bring additional internal or external economies? Or is the industry already so large as to possess practically all the advantages that mere size can give?

In industries of decreasing cost it is clear that the average costs of the typical firm must decrease as the industry expands output. In part, this may be because the average firm is able to approach closer to its optimum size, in part because it shares in the external economies that result from expansion of the entire industry. In industries of constant cost, the costs of the typical firm remain constant; in industries of increasing cost, they rise. In the latter case the rise in the costs of the typical firm is not a result of expansion beyond optimum size. Rather it results from the fact that the total supply of one or more factors of production is limited, so that when the output of the entire industry is increased, the principle of diminishing returns comes into operation.

Relation of long-run cost trends to price. In the short run an increase in the demand for a product will always cause an increase in price. This will also be true in the intermediate period. In the long run, however, the effect of an increase in demand will depend on the cost conditions to which an industry is subject. In the case of industries of increasing cost, an increase in the demand for a product will mean a higher price even in the long run. With industries of constant cost, however, an increase in demand will have no effect on the long-run equilibrium price; in the case of industries of decreasing cost, an increase in demand will, in the long run, actually cause the price to fall. These three long-run price situations are illustrated in Figs. 2, 3, and 4.

In Fig. 2 *DD* represents the original demand for the product and *SS* is the long-run supply or cost curve. Since *SS* slopes upward to the right, it indicates that, even in the long run, this industry can increase produc-

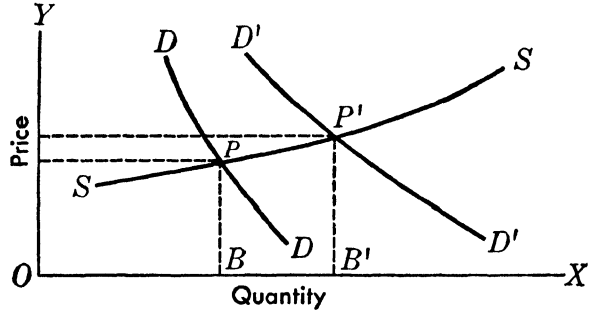


Fig. 2.—Long-run price under increasing cost.

tion only by increasing costs. Consequently, to induce producers to expand, the price must rise. The original long-run equilibrium price is represented by *PB*. *D'D'* represents an increased demand and *P'B'* the new long-run equilibrium price.

In Fig. 3, *DD* represents the original demand and SS the long-run supply or cost curve. *PB* represents the long-run equilibrium price. In

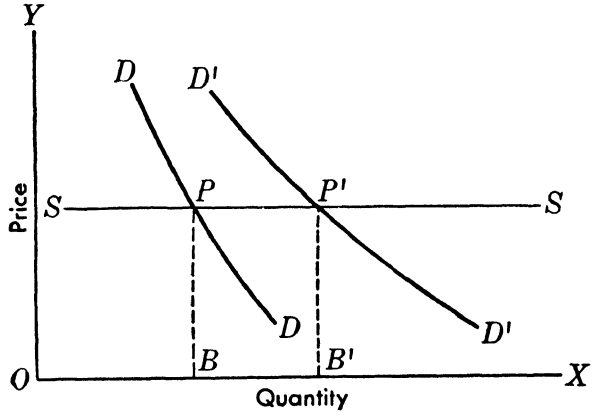


Fig. 3.—Long-run price under constant cost.

this case, however, the cost curve SS is horizontal, indicating that the industry can expand production in the long run without affecting the average cost of the product. An increase in demand will, of course, raise

the short-run and intermediate-period price,[1] but when long-run equilibrium is finally reestablished, the price will be the same as before. $D'D'$ represents an increased demand, and $P'B'$ the new long-run equilibrium price. However, the new price $P'B'$ is equal to the old price PB.

In Fig. 4 DD again represents the original demand and SS the long-run supply or cost curve. PB is the long-run equilibrium price. But in this case the cost curve, SS, slopes downward to the right, indicating that long-run expansion by the industry will actually reduce average unit

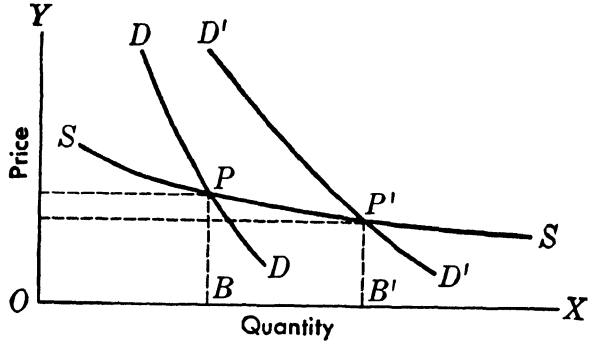


Fig. 4.—Long-run price under decreasing cost.

costs. Here again, an increase in demand would raise the short-run or the intermediate-period price; but when long-run equilibrium was finally established, the price would be lower than before. $D'D'$ represents an increased demand, and $P'B'$ the new long-run equilibrium price.

Summary

This chapter has explained the determination of price in the long run under conditions of perfect competition. It was found that the price of a commodity tends to equal average unit costs for the industry as a whole. Those producers, who have average unit costs above the price, in the long run will go out of business or contract their plant capacity. The result is that the supply of the commodity will decline and the price will rise. On the other hand, producers, who have average costs below the price, in the long run will expand plant capacity and production. If most firms are in this favorable situation, new producers will enter the industry. As a result, supply will increase and price will

[1] See Fig. 5, p. 142.

fall. The price of a commodity is therefore in long-run equilibrium only when it coincides with average unit costs for the industry as a whole. When the price of a commodity is in equilibrium for the long run, it is also in equilibrium for the intermediate period.

In the long run not only does the price of a commodity tend to equal the average unit costs of the industry, but there is also a tendency for all firms in the same industry to have about the same costs. However, this tendency would not completely work itself out even under perfect competition.

In some cases, expansion of production by a whole industry will cause costs to increase in the long run; in other cases, it will not affect them; in still other cases, it will cause them to fall. Since in the long run the price of a commodity tends to coincide with average cost, when an industry is subject to *increasing cost* an increase in demand and in production will raise the long-run equilibrium price. In industries subject to *constant cost,* an increase in demand will have no effect on long-run price; in industries subject to *decreasing cost,* an increase in demand will actually cause the long-run price to fall.

In the next chapter we shall drop the assumption of static conditions in order to study the process of price determination under conditions that are changing, or *dynamic.*

Study Questions

1. Contrast conditions in the long run and in the intermediate period.
2. Why is average cost so significant in the long run?
3. Explain carefully why the average costs of all firms in the same industry tend to equalize in the long run.
4. What relationship is found between average costs in the intermediate and in the long-run period?
5. Can any system be completely static? Why or why not?
6. What conditions cause an industry to become a constant-cost, increasing-cost, or decreasing-cost industry?
7. Show the relationship of price to cost in each of the long-run cost trends mentioned above.
8. If all factors of production would tend to receive the same return no matter where used, would this cause all industries, even though they manufactured different commodities, to have the same average costs? Explain carefully.

9. In the long run under perfect competition if any one producer finds that his average costs are lower than the average of the others, what should he do?

10. Are industries that require a very large investment in plant and durable equipment necessarily industries of decreasing cost? Discuss.

11. A certain commodity is subject to conditions of both competition and decreasing cost. How, in the long run, would an increase in demand affect its price?

12. Why is farming a good illustration of an industry of increasing costs?

Exercises

1. A few years ago most cigars were handmade. Cigar manufacturers used very little equipment and generally operated in low-rent areas. Skill in making cigars was not an important factor. In fact the supply of cigar makers was generally sufficient to make wages low in the industry. Suppose that in this industry the price of materials used was 2 cents per cigar, that labor was 2 cents, and that overhead costs were ½ cent.

a. Under the conditions that prevailed in this industry, what type of long-run costs would you expect to find?

b. At about what price would you expect the product to be sold?

c. Would you expect this industry to operate on a large or small scale? Why?

2. It is often maintained that the street-railway industry eventually becomes an industry of increasing costs. Would such factors as increasing the area of the city, an increase in the amount of service given per rider because of the growth of suburban areas, and competition with other means of transportation have any effect upon cost trends in this industry? Explain carefully.

3. Suppose five electric plants were operating under fairly comparable conditions. Plant 1 has a capacity of 1 million units and, when operating at capacity, the average cost per unit of electricity produced is 7 cents. Plant 2 with a capacity of 2 million units finds its average cost at capacity to be 5 cents. Plant 3 with a capacity of 4 million units has average capacity costs of 3 cents. Plant 4 with a capacity of 7 million units has average capacity costs of 2½ cents. Plant 5 with a capacity of 8 million units has average costs of 2.7 cents.

a. Diagram this condition and point out whether this is an industry of decreasing, constant, or increasing costs.

b. Under the conditions stated what size of plant would you consider the optimum for this industry? Why?

c. What would you expect to be the long-run costs in this industry?

4. Study any particular industry and point out whether you consider it to be an industry of increasing, decreasing, or constant costs. Explain carefully the factors that you took into account in making your decision.

5. Suppose the following conditions prevailed in the long run. What conclusions would you draw as to long-run cost trends in the industry?

a. The industry requires huge amounts of capital.

b. Small amounts of capital are needed but large numbers of unskilled laborers are required.

c. Overheads in the industry are small; little skill is required.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 14.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Part I, Chap. 7.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 10.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 11.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 10.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chaps. 19 and 20.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 16.

Marshall, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book V, Chap. 5.

Taussig, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 14.

15. Price under Dynamic Conditions

In the preceding chapters most of the analysis of price determination has been based on two broad assumptions, namely, that competition was perfect and that the general conditions underlying economic activity remained static. Both of these assumptions are contrary to the facts of the real world of business, but it was necessary to make them in order to present certain basic principles of which the businessman should have a clear understanding. By temporarily disregarding some of the factors that operate in the real world, it was possible to present a clearer picture of others.

Throughout most of the present chapter the assumption of conditions approaching perfect competition will be continued; however, the assumption of a static economy will be dropped. This will bring us a step closer to reality. In the actual world, conditions are dynamic or changing. As a result, prices are constantly thrown out of equilibrium so that new adjustments must be made. The present chapter will show some of the ways in which the dynamic factors in the economy make it necessary to modify the explanation of price determination that has already been presented.

Even in a static economy, in order to show that all prices would eventually come into complete and stable equilibrium it is necessary to make some rather extreme assumptions. For example, in the case of agricultural products it is necessary to assume that weather conditions are the same season after season. However, if the required assumptions are made, it can be shown that equilibrium will finally be established, not only for the short run and the intermediate period, but also for the long run.

In a dynamic economy the situation is quite different. There change is

continual, and as a result price equilibrium is never completely achieved. Before adjustment to one set of circumstances is finished, conditions have changed and the equilibrium prices themselves have shifted. Nevertheless, long-run equilibrium prices are important even in a dynamic economy where they are never actually attained. They are important because they represent the points toward which actual prices are being pulled by the underlying conditions that exist in a market at any given time.

Actual economic conditions are always dynamic, but the rate at which changes take place, as well as their importance, varies greatly. In some countries and in some periods change is relatively slow and regular, so that the price adjustments that take place are similar to those found in a static economy; in other countries or at other times change is relatively rapid and unpredictable, so that long-run adjustments never even approach completion. There are also great differences between industries. In some, conditions are relatively static; in others, they change rapidly and in ways that cannot be foreseen.

Price determination under dynamic conditions. Any change in supply or demand conditions will disturb price equilibrium unless its effect is exactly offset by some other change. Some changes will have only short-run effects; others will operate only in the intermediate period or the long run; and still others will change equilibrium prices for all three periods. Once a change has occurred in conditions controlling the supply or demand for a product, new forces are set in motion; and these new forces immediately begin to pull the price toward new points of equilibrium.

In a highly competitive market the point of short-run equilibrium may be reached or nearly reached very quickly. Suppose, for example, that a great storm should destroy a substantial part of the wheat crop of the country. As soon as the news reached the Chicago Board of Trade, holders of wheat would raise the prices at which they would be willing to dispose of their stocks; and the market price of wheat would rise immediately, let us say by 50 cents a bushel, and at that point be in temporary equilibrium. It is not likely, of course, that such an equilibrium would hold for very long, but in a few days a fairly stable short-run or market-period equilibrium might be established.

Though price adjustments for the short-run or market period may be made quickly, adjustments for the intermediate period and the long run

take considerable time. Suppose, for example, that an artificial fiber like nylon is developed and found to be a good substitute for silk. The immediate effect on the price of silk may be slight. There will be small inducement for present holders of stocks of silk to sell it at lower prices, because it will be months or years before factories can be built and the new product brought to the market in appreciable quantities; but in time, use of the new product may greatly reduce the demand for silk and so bring about a great reduction in its price. However, in a dynamic economy long-run price adjustments are never entirely completed. As has already been noted, before equilibrium is achieved under any given set of conditions, changes are sure to occur which will move the equilibrium points themselves to new positions.

The reader will better understand the process of price determination in a dynamic economy if he compares it with the process in a static economy. It has been pointed out that in a static economy, under sufficiently rigid assumptions, the price of a commodity will eventually reach stable equilibrium for the short run, the intermediate period, and the long run. Moreover, the equilibrium price will then be the same for all three of these periods. This situation is illustrated by Fig. 1.[1] SS is the long-run supply curve and also the long-run average cost curve for the industry; S_1S_1 is the short-run supply curve; [2] and S_2S_2 is the intermediate-period supply curve. DD is the demand curve; and since all three supply curves intersect DD at P, PR is the equilibrium price for the short run, the intermediate period, and the long run. PR is also the actual market price and, therefore, it may be said that price is in equilibrium for all three periods. It should be understood that when price is in long-run equilibrium it is always in equilibrium for the short run and the intermediate period.

Let us now suppose that conditions change; for example, let us suppose that there is a rapid and substantial increase in demand which gives every indication of being rather permanent. This change in demand will shift the equilibrium prices for all three periods. What is more important, it will cause them to diverge from the actual market price and to sepa-

[1] Before going further the reader should review Fig. 5, p. 142.

[2] Since the curve S_1S_1 is vertical, it indicates that in the short run the quantity of this commodity offered for sale cannot be varied. For most commodities this would not be strictly true.

rate from one another. No longer will the equilibrium price be the same for the short run, the intermediate period, and the long run. This separation of the equilibrium prices for the three periods is typical of a dynamic economy. Although in a dynamic economy the actual market price is always being pulled toward the point of short-run equilibrium,

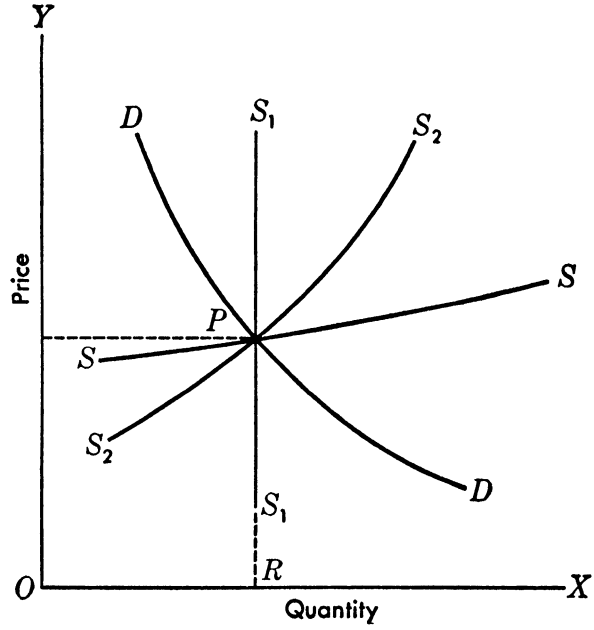


Fig. 1.—Equilibrium in the short run, the intermediate period, and the long run.

and this toward the point of intermediate-period equilibrium, and this in turn toward the point of long-run equilibrium, the process is never completed before new changes intervene. Figure 2 illustrates the separation of equilibrium prices in a dynamic economy. The increased demand is represented by D_1D_1. As a result of this change in demand, the short-run equilibrium price has risen from PR to P_1R; the intermediate-period equilibrium price has shifted from PR to P_2K; and the long-run equilibrium price has shifted from PR to P_3L. Because short-run adjustments are made quickly, the actual market price will very soon approximate the short-run equilibrium price, P_1R. Then the forces of the market will begin pulling the short-run price, P_1R, and later the intermediate-period price, P_2K, toward the long-run price, P_3L.

If the market referred to by Fig. 2 were static instead of dynamic, in time P_1R, P_2K, and P_3L would coincide. This would come about in the following way: The high short-run price P_1R would cause producers to expand output in the intermediate period in the plants already operating. As the current rate of production rose, the short-run supply curve, S_1S_1, would, of course, move to the right. However, an increase in the output

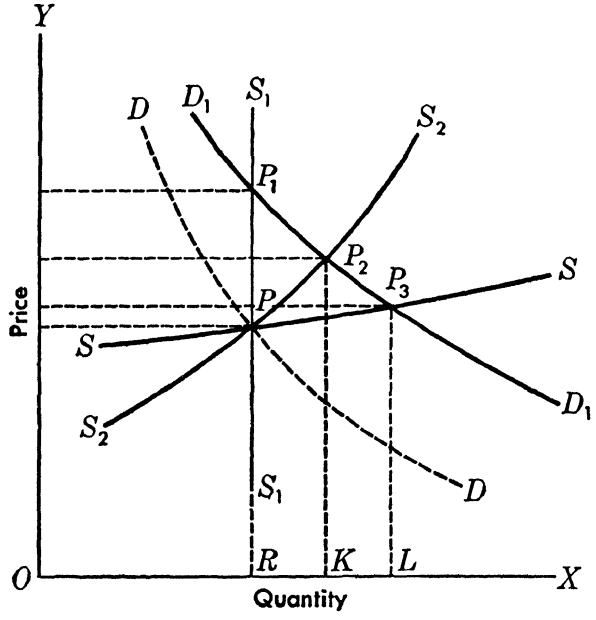


Fig. 2.—Separation of equilibrium prices in a dynamic economy.

of existing plants would be at the expense of sharply rising costs; and to avoid these higher costs plant expansion would be undertaken. As soon as new plants were built and came into operation, the intermediate-period supply curve, S_2S_2, would also begin to move to the right. Under static conditions eventually S_1S_1 and S_2S_2, as well as SS, would all pass through the point P_3; equilibrium price would again be the same for all three periods of time; and actual price would once more coincide with equilibrium price. The situation would again be similar to that shown in Fig. 1.

Since, however, Fig. 2 refers to a market that is dynamic and not static, the kind of complete price equilibrium just described is never reached.

Before one set of adjustments is completed, either supply or demand changes again, and new adjustments begin. There are many factors that affect supply and demand, and under dynamic conditions some of them are always changing in ways that cannot be predicted. As a result, the short-run, intermediate-period, and long-run equilibrium prices often differ very widely; and even under the most favorable conditions they are never exactly the same.

In spite of this variation in the factors controlling supply and demand, the reader should not conclude that it is entirely useless to try to predict price trends under dynamic conditions. It is true that such predictions are always subject to a margin of error. Nevertheless, the businessman who really knows his market and who also understands economic principles can sometimes forecast prices with surprising accuracy. This is especially true in those cases where the price of a commodity is largely controlled by one or two basic factors, and where these factors are changing slowly and at a relatively constant rate.

Business adjustments in a dynamic market. Short-run and intermediate-period considerations are always pressing upon the businessman, but he must also take long-run factors into account. Business is operated in the present on the basis of forecasting the future. The air transportation industry, for example, may have plant capacity today which in the light of today's demand seems ridiculously large; yet persons charged with the operation of the business must consider the future, and in the light of the future the capacity may not be too great. Management in the telephone industry maintains that, in order to operate efficiently and be ready for all demands, it must estimate probable needs 40 to 50 years in the future and must act today in accordance with the costs involved. Thus, an industry may appear to be maladjusted so far as the market is concerned in the present or near future; and yet in the long run its operations may prove to have been very wise. Consequently, when the forces affecting prices are being discussed, we must always remember that long-run as well as short-run or intermediate-period considerations are operating. However, in a dynamic economy the nature of change is to a large extent unpredictable. Although it is necessary to plan for the future, such planning is useful only up to a certain point; and the longer the period for which plans are made, the greater is the likelihood of making wide errors of judgment.

Representative firm. In the preceding chapter it was explained that under perfect competition and static conditions the long-run price of a commodity would equal average costs for the industry as a whole. However, the assumption of static conditions was not pushed so far as to exclude changes in the position of the individual firm. Even in a static economy the conditions affecting individual firms would never be entirely static because life would have to go on, and therefore each firm would be subject to unpredictable changes in the efficiency and personnel of its management. As a result, the costs of individual firms would always be changing, and no matter how long a period of time might elapse, in any competitive industry there would always be some firms making profits, others taking losses, and still others just breaking even. Price would not be equal to the average costs of each and every firm, but rather to the costs of an average or *representative* firm.

The representative-firm concept was first developed by a noted English economist, Professor Alfred Marshall, who developed this concept in order to explain, in the first instance, the relationship of price to the costs of the individual firm in a "stationary state," that is to say, in a static economy.[3]

Representative firm under dynamic conditions. Although Marshall first applied the representative-firm concept to a static economy, it can also be employed to explain the relation between long-run price and the costs of the individual firm under dynamic conditions, especially in industries where there are a large number of competing producers.[4] Since in a dynamic market conditions are constantly changing, there may be years in which most firms are making profits. Nevertheless, if a given average rate of production is to be maintained over the long run, price in normal years must be high enough to cover the economic costs of the representative firm.

As conceived by Marshall, the representative firm is not the most efficient nor yet the least efficient in the industry, but a firm whose average costs are just low enough so that it can continue to operate in the market indefinitely.[5] Thus the representative firm is not strictly a

[3] MARSHALL, ALFRED, "Principles of Economics," 8th ed., p. 367.

[4] Marshall does not use the expression "perfect competition," but he apparently had in mind a situation more or less approaching the one now commonly described by this term.

[5] MARSHALL, *op. cit.*, p. 317.

mathematical average, but an average in the sense that it is a typical producer, whose costs are neither high nor yet unusually low. Figure 3 represents the relation of long-run price to the average costs of the representative firm.

In the market referred to by Fig. 3 a large number of firms is assumed to be producing and selling a certain commodity. For the sake of simplifying the illustration only six are shown, A, B, C, D, E, and F;

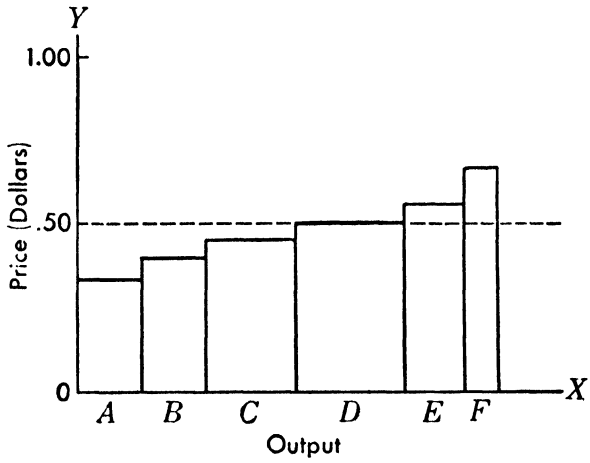


Fig. 3.—Relation of the average cost of the representative firm to long-run price.

but each of these may be thought of as representing an entire group. The figure shows that sellers are producing a sufficient quantity of goods to cause demand and supply to equate in the long run at a price of 50 cents per unit. This price is just equal to the average costs of the representative firm, D. Firms A, B, and C have costs below the price and are making substantial economic profits. It is clear, however, that the firms on the fringe, represented by E and F, are operating at a loss.

However, high-cost firms that produce at a loss are not immediately removed from the market. In fact, there are always some firms in these high-cost positions. Sometimes such firms have high costs because they have just begun operation and have not yet reached an efficient size. Frequently certain producers have high costs because of inefficient management or obsolete machinery and equipment. Thus, although they are on the way out, they have not yet reached the point where they are forced to stop production. At present prices they can still cover their

operating costs and have a small amount left for overhead; but unless they increase their efficiency, they will ultimately be forced out of business. In a dynamic economy some of them may, of course, be saved by a change in market conditions that will raise the price of the product. However, even if some firms are forced out of business, their places are taken by others. For reasons that will shortly be explained, the group of high-cost producers continues to exist.

Representative firms in actual competitive markets. Marshall recognized that the representative firms in an industry are not necessarily the same from year to year. New firms are constantly being organized. Some are successful from the start, but others encounter unexpected difficulties. Moreover, the costs of individual firms continually change. Some low-cost firms lose in efficiency and their costs rise. High-cost firms eventually retire from business or else move into the low-cost brackets. Professor Secrist of Northwestern University made statistical studies which prove that such shifts are not just theoretical but that they really occur in the actual business world. Moreover, he found that the change in costs is greatest at each extreme. High-cost firms lower their costs perceptibly if they remain in business; whereas low-cost firms, either because by increasing average costs they can increase profits or because they are subjected to many types of pressure from management, labor, and even customers for a greater share in the proceeds, soon operate under higher average costs.

The findings of Secrist are illustrated by Fig. 4, which shows possible average cost trends over a period of years, as between six individual firms, all producing a certain commodity and selling it in competition with each other. The chart shows that in the first year A is a high-cost firm, while F is the lowest cost firm. Then as the years pass, average costs of A decrease. This may be due to an increase in volume of business or a more efficient organization. Costs of firm F increase, while the firms in between have varying cost trends, some having increasing costs and some decreasing. At the end of 7 years E, which was a fairly low-cost firm, has become the high-cost firm, while B has become the lowest cost firm. Perhaps if the period were extended, B's cost of production would again start upward while E's would start downward. The point to observe is that the relative positions of firms change from year to year, but differential costs persist.

We should observe, however, that there is a constant pull in costs of production to a point that is somewhat lower than the average of the highest cost producer but higher than that of the most efficient producer.

Bulk-line cost. Closely related to the representative-firm idea is the concept of bulk-line cost. The bulk-line cost concept grew out of an attempt to answer the question, What price must be paid to induce producers in a competitive industry to supply a given amount of product? It is obvious that, if the price offered is only just equal to the cost of

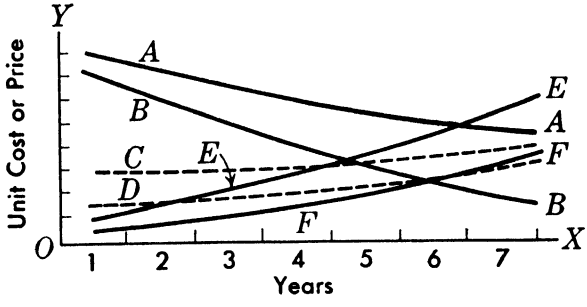


Fig. 4.—Long-run cost trends of individual firms.

production of the lowest cost producer, all other producers will suffer a loss, and hence they will be inclined to contract output or even to withdraw from the market completely. If, on the other hand, a price is offered that equals the cost of production of the highest cost producer, some firms will make excessive profits. Moreover, statistical studies like those of Secrist indicate that to maintain a given rate of production in a competitive industry it is not necessary for the price to cover the costs of the least efficient firms.

During the First World War attempts were made to secure goods for the government at fair prices; those in power took as a starting point for their thinking the concept of the representative firm. They then changed from a study of the costs of any one firm to a study of the price necessary to bring out the quantity desired. They sought first to answer the following question, What would be the minimum cost of producing various percentages of the goods required? That is, how much would it cost to produce 10, 20, 30, and so on, up to 100 per cent of the total amount desired? Knowing, however, that in competitive industries price need not equal the cost of production of the highest cost firm, they tried to

determine the portion of the product whose costs must be covered by the price. They were interested in getting a certain total amount of product while not allowing low-cost producers to make exorbitant profits.

It was found that, if the average cost of securing the whole quantity sought for was used as a base, as a rule approximately 70 per cent of the total amount could be produced at a cost that did not vary more than

Table 14.—*Average Costs, by Groups, of the Firms Producing a Certain Commodity*

Group	*Output*		*Average unit cost*	*Per cent of total output*	
	For each group	*Cumulative totals*		*For each group*	*Cumulative totals*
A	50,000	50,000	$1.00	25	25
B	44,000	94,000	1.25	22	47
C	40,000	134,000	1.30	20	67
D	36,000	170,000	1.35	18	85
E	30,000	200,000	1.40	15	100

20 per cent either above or below this average cost. It was further found that about 15 per cent of the commodity could be produced at a cost that was even below this range. In other words, about 85 per cent of the quantity demanded could be produced at costs none of which was more than 20 per cent higher than the average cost for the total output. They then found the average cost of production of the highest cost firm whose output was necessary to produce this 85 per cent of the product; and they considered that cost to be a fair price for the entire output. An example may aid in making this clear.

Suppose that in a certain industry there is a large number of competing firms. Also, to simplify our illustration, suppose that these firms fall into five groups and that the costs and outputs of these groups are as represented in Table 14, which shows that the cumulative or total output of these five groups of firms is 200,000 units. Suppose that the govern-

ment wishes to acquire just this amount of the commodity. The question is, How much must it pay in order to secure the 200,000 units? If our theory of the relation between price, bulk-line cost, and volume of production is correct, price need cover the cost of producing only 85 per cent of the output needed. Since groups A, B, C, and D produce 85 per cent of the amount required, if the price is high enough to cover the average costs of production of the firms in group D, the desired total output will be forthcoming. Consequently, a price of $1.35 is offered for the good, since that amount equals the cost of the bulk of the output. This price is called the *bulk-line* price.

Let us consider for a moment what is likely to take place when a price of $1.35 is offered. Will production at once drop to 1,700 units? This is unlikely, for even though the firms in group E are not securing a price equal to their lowest average costs of production, they will continue to produce until either they fail or succeed in reducing their costs, or are able to shift into some other industry. In the meantime some low-cost firms may expand and new firms will enter the market; and this will tend to prevent any contraction in total output. Since bulk-line price is above the average costs of production of the industry as a whole, it serves to attract new firms whose owners believe they will be able to operate efficiently. However, the E group of high-cost firms will never disappear for, though some firms move into lower cost groups and some fail, others are constantly taking their places.

The theory of bulk-line price may be summed up as follows: (1) It is based on the representative-firm concept. (2) It is also based on statistical studies, the findings of which may vary somewhat for different industries, or for the same industry if conditions change. (3) It applies to actual situations in which conditions are dynamic and competition is not entirely perfect. (4) It applies, however, only to industries in which competition is free and effective, that is to say, to industries in which there are a large number of producers who are actively competing. (5) It attempts to determine the price necessary to maintain a given volume of output of a commodity. (6) It holds that, in a typical competitive industry, price in the long run must equal the average cost of production of the firm whose product is necessary to bring about 85 per cent of the total output of the commodity to the market.

Summary

The purpose of this chapter is to explain the determination of prices under dynamic conditions. Under static conditions and perfect competition the short-run, intermediate-period, and long-run equilibrium prices of a commodity eventually coincide with one another and with the actual price at a point of stable equilibrium. In a dynamic economy, however, the equilibrium prices for the three periods of time are never the same. This is because supply and demand conditions keep changing. The actual price at any given time is likely to be close to the short-run equilibrium price, but it will seldom agree closely with the intermediate period or the long-run point of equilibrium. However, intermediate-period and long-run equilibrium prices are important in a dynamic economy because they represent points toward which, at any given time, actual prices are being pulled; and even in a dynamic economy, if basic changes in supply and demand conditions are not too rapid, the major intermediate-period and long-run adjustments may be largely completed.

In a dynamic economy the concept of the *representative* firm is an aid to understanding the relationship between long-run competitive prices and the costs of individual firms. It was found that the long-run price of a commodity tended to equal, not the average costs of each firm in the industry, but the average costs of a typical or representative firm. In a competitive industry there will, even in the long run, always be firms making profits, firms breaking even, and firms taking losses. One might expect the high-cost firms to be forced out of business or move into a lower cost group. This does in fact take place. However, the high-cost firms as a group do not disappear, because as some firms move out of the group, others take their places.

Another concept useful in understanding the relation of price to cost under the dynamic conditions found in the actual business world is *bulk-line cost.* The theory of bulk-line cost can best be applied to industries where, though competition is not entirely perfect, there are a large number of competing firms. The theory of bulk-line cost is based on the representative-firm concept. It attempts to determine, on the basis of statistical studies, the price necessary to maintain a given volume of output in a competitive industry. According to this theory, price in the long run must equal the average cost of production of the firm whose product is just necessary to bring 85 per cent of the total output of the commodity to the market.

From our study of price under dynamic conditions we conclude that, where competition is prevalent, prices tend to coincide with the cost of production of the representative producer, or tend to coincide with bulk-line cost.

Study Questions

1. In what basic assumption does this chapter differ from those preceding it?
2. Define the terms "long run," "intermediate period," "dynamic market," and "normal price."
3. How is the long-run cost price relationship determined in a dynamic competitive market?
4. What is the relationship between marginal costs and price in the long run? Why is this relationship significant?
5. Why do we say that in a dynamic economy long-run price adjustments are never entirely completed?
6. Is the concept "representative firm" significant today? Why or why not?
7. How would you determine the bulk-line cost in any industry?
8. In a dynamic market why is it necessary to consider cost conditions that prevail in individual firms as well as those cost conditions which prevail in the industry as a whole?
9. In the light of the discussion, just what cost-price relationship tends to prevail in the long run under dynamic conditions?

Exercises

1. Some economists maintain that since the pull of competition tends to force all production costs into very narrow variations, it is not necessary to consider dynamic conditions.

a. Justify their point of view.

b. Refute it.

2. In a certain industry 26 per cent of the product can be produced at an average unit cost of 7 cents. However, an additional 24 per cent can be produced only at an average unit cost for this addition of 9 cents; another 36 per cent can be produced at an average unit cost of 10 cents, while another 14 per cent can be produced at an average unit cost of 14 cents.

a. Draw a diagram illustrating the above condition under the assumption that the total production is 100 units.

b. In the long run if all costs remain constant and if the demand is for 100 units, what price will prevail in the market?

c. What would be the bulk-line price under conditions stated in the problem? Why?

d. What would be the marginal cost of each of the producers listed above if conditions prevailed as in (*b*)? Why? As in (*c*)? Why?

3. Suppose that the government wishes to aid certain business groups by placing a floor under prices. How would the government determine whether the price suggested was or was not fair?

For example, suppose, in the case given above, that the government tried to establish a fair price. What would this price be, and why?

References for Further Reading

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 9.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chaps. 13 and 14.

Marshall, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book IV, Chap. 13; Book V, Chaps. 3 and 12.

Secrist, H.: Bureau of Business Research, *Northwestern University Bulletin* 8, Section II, Bulletin 9.

16. Price under Monopoly

The preceding chapters have explained how prices are determined when many sellers are competing in the same market. The present chapter will explain the determination of prices in those situations where a single producer has complete control over the supply of a commodity, so that, within limits, he can raise the price if he is willing to decrease his sales. To have complete control over supply a producer must, of course, be a monopolist, that is to say, the only seller of the commodity in question. Cases in which one producer has a complete monopoly of an important commodity or service are not very common outside of the utility field; and in that field the government usually steps in to limit his power over prices. Nevertheless, it is important to study the determination of price under complete monopoly for two reasons: (1) Minor monopolies, like those under patents and copyrights, are fairly common. (2) If we understand how price is determined under complete monopoly, as well as under perfect competition, we shall be better able to understand how price is determined in the great number of situations that fall between these two extremes.

There is one difference between the theory of monopoly price and the theory of competitive price that should always be kept in mind. Under perfect competition, price is determined by the impersonal forces of supply and demand. As a result, once the conditions of supply and demand are known, the price is known. Under monopoly, however, price is fixed by the monopolist; and no theory that can be developed will indicate with certainty the point at which he will choose to set it. The only way to develop a theory of monopoly price is to assume that (1) the monopolist can estimate his costs and the demand for his product and that (2) he will charge the price that will maximize his total profits.

Monopoly is often thought of as the direct opposite of competition, and in a sense this is true. However, even a complete monopoly is not entirely free from any kind of competition. It is true that, if one has a

complete monopoly of a commodity or a service, he has no competitors who are selling the same thing; but he still must meet the competition of substitutes. If the substitutes are poor, he need have little worry. If, however, they are good, he may face the stiffest kind of competition. There is no commodity or service that does not have some competition from substitutes. Buses, elevated lines, and private automobiles compete with streetcars; the postoffice, the telegraph, and all other means of communication compete to some degree with the telephone; and iron, copper, and magnesium alloys compete with aluminum. These are only examples.

There are also other factors that may make it difficult for a monopolist to maintain control over the price of his product. The competition of substitutes has been mentioned, but this is not his only problem. Often he is subject to potential competition, that is to say, to the possibility that someone else will begin to produce the same commodity that he makes. In some cases, of course, he is protected from direct competition by patents or copyrights. At best, however, these give him secure control of his market for only a limited period of time. Moreover, even if his monopoly position seems secure, he will often be wise to exercise his power over price with great discretion. If the commodity that he produces is important and if the public believes that he is charging an unreasonably high price, the government may step in and set a price that allows him only a moderate return. In this chapter, however, we are assuming an extreme case, namely, that the monopolist has complete control of supply and is free to set the price that will give him the greatest total profit.

Fixed stock. The simplest example of complete monopoly is found where the seller is in possession of the entire stock of some commodity that cannot be reproduced. For example, the monopolist may own a painting by a noted artist, the only one of its kind in existence. Under these conditions, he will seek to secure the highest possible price, for the more he receives, the higher will be his profit. Figure 1 shows that, if the monopolist were to receive a price *PB* for his good, his total income would be represented by the rectangle *APBO*. On the other hand, if he received a price of *P''B*, his total income would be *A''P''BO*. Moreover, since the stock is fixed, total costs are fixed; hence, the higher the price, the larger his total and also his net profit.

Of course, if the monopolist in this illustration had several paintings that were alike, it might pay him to set a price low enough to sell two or three instead of only one. His aim should be, not to sell a single picture at the highest price possible, but rather to receive the largest possible sum for his entire stock of paintings. Since his costs are fixed, this would also give him the highest total as well as the highest net profit; or, if a profit could not be made, it would reduce his loss to a minimum.

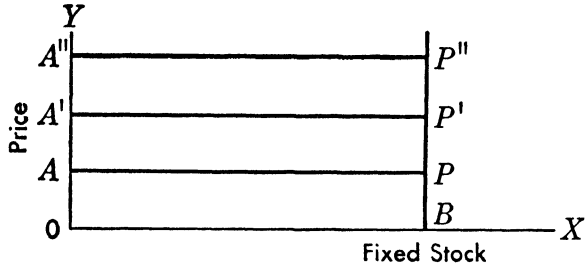


Fig. 1.—Total income of monopolist at various prices for his product.

Reproducible stock. Monopoly is also found where the stock of a good is reproducible. Producers of certain commodities are constantly attempting to secure monopoly by making agreements among themselves. For this reason cartels and various so-called valorization and rationalization schemes have been used by businessmen and by governments to secure monopoly profits. Thus nitrates, coffee, rubber, tin, camphor, and many other commodities have been monopolized. In the past in the United States great trusts arose in such products as oil, tobacco, steel, whisky, and sugar. Moreover, there exist today natural monopolies, such as public utilities. These, as we have noted, are regulated by the state.

When the monopolist is in a position to control a reproducible good, he must consider both his cost of production and the demand for his product. The monopolist is interested in securing the highest total net profit, not the highest gross sales income. For this reason, trends in cost as well as in demand must be carefully considered.

Average revenue. The monopolist knows that the price that he can charge for his product is dependent upon the demand schedule and that, in order to charge a little more, he must be willing to sell a little less. The exact result will depend upon the elasticity of the demand for his

good. Thus, unlike the producer under competition whose demand may be represented by a horizontal line, the monopolist finds that he can regulate the price at which the commodity is to be sold by exerting control over the supply. The total price for which a commodity is sold is the seller's gross sales income, or total revenue.[1] To the monopolist the demand curve is an *average revenue* curve. If we assume that all units of a good must be sold at the same price, the demand curve shows him what his average revenue per unit will be if he places various quantities of a good on the market.

Marginal revenue. Obviously the principal difference between a monopolist and a producer in a competitive market is that the monopolist can fix the price of the product, while the producer subject to perfect competition cannot. In considering whether to expand production the competing producer need take into account only the way in which changes in his costs will affect his profits. The monopolist, however, when he contemplates expansion, must consider not only the effect on his costs but also the effect on his selling price and his total revenue. The monopolist knows that he must reduce the price at which he sells his commodity if he wishes to expand production and sales. If to increase sales by a given amount he has to reduce price substantially, his total revenue may actually decrease. At best his revenue will not increase in proportion to the increase in sales. If gross sales income does increase, the addition to total revenue which results from the sale of an additional unit of the product is called *marginal revenue.*

Monopoly price in the short run. Under monopoly, just as under competition, cost of production has less influence on price in the short run than in the intermediate period or the long run. In the short run even the stock of a reproducible commodity is for the time being fixed. If the commodity is highly perishable, we have the case of fixed supply, already referred to. In this situation any portion of the good that is not sold promptly becomes a complete loss. Under competition each holder would be wise to dispose of his entire stock regardless of price. Under monopoly, however, it may be possible to obtain a larger total revenue by selling only part of the stock and letting the rest spoil. The aim of the monopolist is to secure the largest total revenue regardless of what his costs may have been. This policy will maximize his profit or, if a profit is impossible, it will minimize his loss.

[1] The reader should take care not to confuse revenue with profit.

However, if the commodity that the monopolist has in stock is durable, in determining his price policy he should take into consideration the cost of reproducing it. After allowing for the cost of holding it, he ordinarily would not sell it for less than the cost of replacement. Indeed, since he is a monopolist, it is likely that he would insist in the short-run period on getting a price somewhat higher than the cost of reproduction, and that he would carry over into the future any part of his stock that he could not sell at that price. A study of the intermediate period will throw more light on the policies that a monopolist will follow.

Monopoly price in the intermediate period. In the intermediate period the monopolist's plant capacity is fixed, but he can expand or contract production by increasing or decreasing his use of the variable factors. Whether he can increase his profits by expanding production will depend on the relation between his marginal cost and his marginal revenue. As long as the increase in total cost that results from producing one more unit of product is less than the increase in total sales revenue that results, it is clear that producing the additional unit increases profits. When the increase in cost and the increase in revenue are equal, production of the unit is a matter of indifference. As soon, however, as the increase in cost is greater than the increase in revenue, producing an additional unit of product decreases profits. Therefore, *to maximize his profits in the intermediate period, the monopolist should keep increasing production until the point is reached at which marginal cost and marginal revenue are equal.*

Let us illustrate this principle by an example. Suppose a monopolist is able to sell 100 men's suits per day at $75 each, thus receiving a total revenue of $7,500. The suits cost him $50 apiece, or a total of $5,000, and his profit per day is $2,500. Since he is interested in the highest possible profits, he decides as an experiment to make and sell 125 suits per day. He finds, however, that if he wishes to sell 125 suits per day he must reduce his price for a suit to $70. This, however, gives him a total revenue of $8,750, or an increase of $1,250. Let us suppose that the additional cost of producing the extra 25 suits is $1,000. If we think of the 25 suits as the unit by which production is increased, it is clear that the marginal revenue of $1,250 exceeds the marginal cost of $1,000 by $250, and that this is the amount added to the monopolist's profit, so that his profits per day are now $2,750. Moreover, since the sale of the additional 25 suits increased his revenue by $1,250, the average marginal

revenue from each of the extra 25 suits is $50. It is also clear that the average marginal cost is $40. Therefore, each of the additional suits increases the monopolist's profits by $10.

The monopolist in the example just given will find that his total net profits will increase as long as he can keep marginal cost below marginal revenue. For example, if the marginal cost of increasing production by

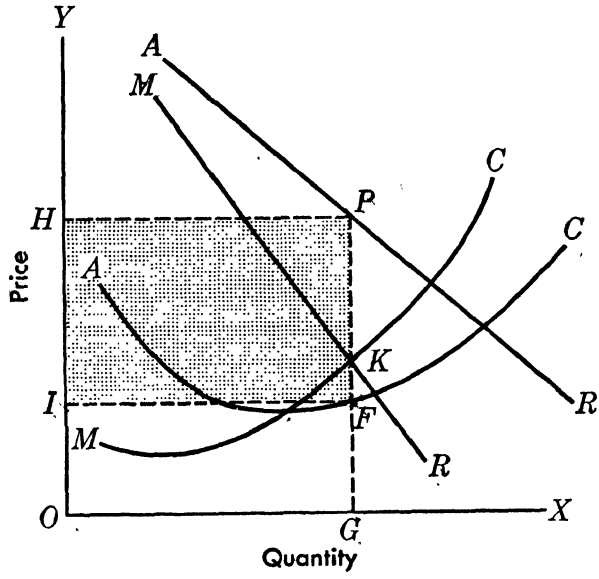


Fig. 2.—Determination of monopoly price in the intermediate period.

yet another 25 suits per day should average $43 per suit and the marginal revenue received should average $46 per suit, there would be a further increase in profits. As soon, however, as marginal cost equals marginal revenue, profits will cease expanding; if marginal cost exceeds marginal revenue, they will actually contract.

Figure 2 represents graphically the determination of monopoly price in the intermediate period. *AR* represents average revenue and *MR* marginal revenue; *AC* represents average cost and *MC* marginal cost. It pays the monopolist best to produce *OG* units of product, because at this point marginal cost equals marginal revenue. This is shown by the intersection of *MC* and *MR* at *K*. When the monopolist produces *OG* units, he can obtain a price equal to *PG*. His average cost is then *FG*

and his average profit *PF*. His total profit is therefore *PF* times *OG*, or *PF* times *IF*. This is represented by the rectangle *HPFI*. Any amount of product greater or less than *OG* will yield a smaller total profit.

It should be noted that the marginal revenue curve, *MR*, falls faster than the average revenue curve, *AR*. This will always be the case when a monopolist sells at a single price. In order to sell an additional unit of product, he must reduce the price of all units, and the increase in his total revenue will therefore be less than the price that he receives for the extra unit.

Monopoly price in the long run. In the long run the monopolist can vary the rate of production not only by increasing or decreasing the output of existing plants; he can also increase or decrease the size and number of his plants. By adjusting his plant capacity he can, in the long run, produce any given volume of output at the lowest possible average unit cost for that amount of output. In a monopolized industry there is, of course, only one firm. If this firm has a single plant it may, by increasing the size of the plant, gain more and more of the advantages of mass production. The same may be true even if it already has several plants, as it may gain advantages by operating the enterprise as a whole on a larger scale. In either case, if expansion is carried far enough, the firm will eventually reach its optimum size, and it will then be able to reduce average long-run unit costs to the lowest possible level. When costs are near this low point, there may be a wide range over which capacity and output can be expanded at approximately constant cost. If, however, the firm expands much beyond its optimum size, long-run average unit costs will begin to rise. The declining phase of a monopolist's *long-run* average cost curve results from the advantages of large-scale production and large-scale organization; while the rising phase of his long-run cost curve is caused by the difficulties of large-scale production and large-scale management after a firm has passed its optimum size.

It is clear, then, that a monopolist, as he expands production in the long-run period, may be subject successively to conditions of decreasing, constant, and increasing cost. In many cases, however, *for any increase in production that is at all possible,* he will be subject to only one of these three cost situations. To simplify our problem, we shall assume that this is the case, and we shall illustrate how monopoly price would be determined in each of these three situations.

If long-run changes in the rate of output change a monopolist's average costs, it is because they change the general efficiency with which his organization operates. The long-run marginal cost curve of a monopolist is *derived from* his average cost curve. When average cost rises as production expands, marginal cost—the addition to total cost required to increase output by one unit—must be above average cost. When average cost is constant, marginal and average cost must be equal. Finally, when average cost is falling, marginal cost must be below average cost.

In the intermediate period it was found that a monopolist should expand production until marginal cost equals marginal revenue. The same principle applies to the long-run period. To make the greatest total profit in this period, he should expand production until long-run marginal cost is equal to marginal revenue.

Monopoly price under constant cost. Table 15 and Fig. 3 represent the determination of long-run monopoly price under conditions of constant cost.

Table 15.—*Long-run Monopoly Price under Constant Cost* *

Price	*Units produced and sold annually*	*Total revenue*	*Average marginal revenue*	*Average cost*	*Long-run marginal cost*	*Total cost*	*Total net profit*
$1.00	4,000	$ 4,000	$ 1.00	$0.25	$0.25	$1,000	$3,000
0.90	7,000	6,300	0.76⅔	0.25	0.25	1,750	4,550
0.80	11,000	8,800	0.62½	0.25	0.25	2,750	6,050
0.70	15,000	10,500	0.42½	0.25	0.25	3,750	6,750
0.60	19,000	11,400	0.22½	0.25	0.25	4,750	6,650
0.50	22,000	11,000	0.13⅓	0.25	0.25	5,500	5,500
0.40	25,000	10,000	−0.33⅓	0.25	0.25	6,250	3,750
0.30	30,000	9,000	−0.20	0.25	0.25	7,500	1,500

* In this table marginal revenue is found in the following manner: When sales increased from 4,000 to 7,000, or an increase of 3,000 units, receipts or revenue increased from $4,000 to $6,300, or an increase of $2,300. Thus the additional 3,000 units sold increased revenue by $2,300, or an average of $0.76⅔ for the additional units. For convenience, we call this increase per unit *average marginal revenue.*

Figure 3 is based on Table 15 and represents the situation of a monopolist who produces under conditions of constant cost. In this case average and marginal costs are always equal. The figure indicates that the monopolist should expand production until his output is about 18,000 units a year. He will then be able to get a price of about 62 cents, and his average cost, marginal cost, and marginal revenue will all be equal

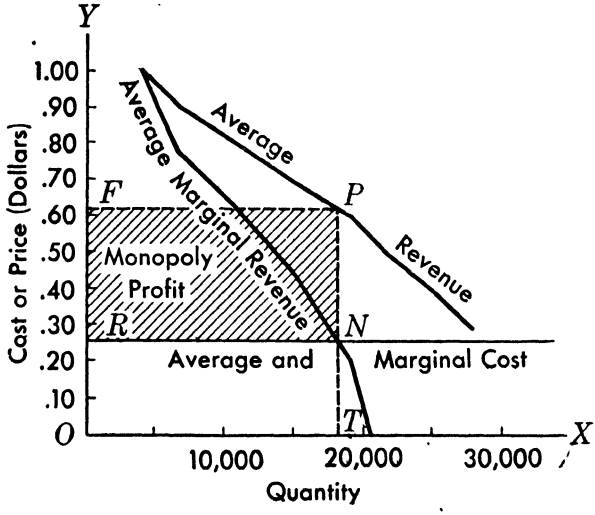


Fig. 3.—Long-run monopoly price under constant cost.

at 25 cents. His average profit per unit will be about 37 cents. Graphically the price is represented by *PT*, the average unit cost by *NT*, and the average unit profit by *PN*. The total profit will be *PN*, the average profit per unit, times *OT*, the number of units produced and sold. This is shown by the rectangle *FPNR*, which represents the maximum annual profits that the monopolist can make in the long run.

We reach similar conclusions by a direct examination of Table 15. There we can see that, if the monopolist insists on fixing a price that is an even multiple of 10, he will make the highest total profit by charging 70 cents. At that price his profit is $6,750, while at 60 cents it is only $6,650. At 70 cents, however, marginal revenue is still slightly above marginal cost, while at 60 cents it is below marginal cost. It is clear, therefore, from the table as well as from the figure, that the price that will yield the maximum monopoly profit is somewhere between 60 and

70 cents since at some point between these prices marginal revenue and marginal cost would coincide.

Monopoly price under increasing cost. Table 16 and Fig. 4 represent the situation of a monopolist who produces in the long run under conditions of increasing cost.

Table 16.—Long-run Monopoly Price under Increasing Cost

Price	*Units produced and sold annually*	*Total revenue*	*Average marginal revenue*	*Total cost*	*Average cost*	*Average marginal cost*	*Total net profit*
$1.00	4,000	$ 4,000	$1.00	$1,000	$0.25	$0.25	$3,000
0.90	8,000	7,200	0.80	2,800	0.35	0.45	4,400
0.80	11,000	8,800	0.533	4,500	0.409	0.567	4,300
0.70	14,000	9,800	0.333	6,800	0.485	0.767	3,000
0.60	18,000	10,800	0.25	10,800	0.60	1.00	0

The chart shows that this particular monopolist, operating under conditions of increasing cost, should in the long run expand production to about 10,000 units per year. We note that, under conditions of increasing cost, marginal cost is always above average cost. When production reaches about 10,000 units, marginal cost and marginal revenue will be equal and both can be represented by the line *KA*. The price will be *PA*, the average cost *NA*, and the average profit *PN*. Total profit will be *PN*, the average profit, times *OA*, the quantity produced and sold. In the chart total profit is represented by the rectangle *FPNR*, and again this is the largest total annual profit that, in the long run, the monopolist can make.

As we should expect, an examination of Table 16 brings us to much the same conclusions as an examination of the chart. We see that, if the monopolist wishes to charge a price that is an even multiple of 10, he will do best to ask 90 cents. At this price his total profit would be $4,400, while the next largest profit, obtained by charging 80 cents, would be only $4,300. However, when he charges 90 cents, marginal revenue is above marginal cost, while at 80 cents it is below. It is clear, therefore,

that to gain the very largest total profit the monopolist must charge a price somewhere between 80 and 90 cents.

Monopoly price under decreasing cost. The monopolist stands in a very favorable position if his long-run average and marginal costs are decreasing. Under these conditions, although every increase in output

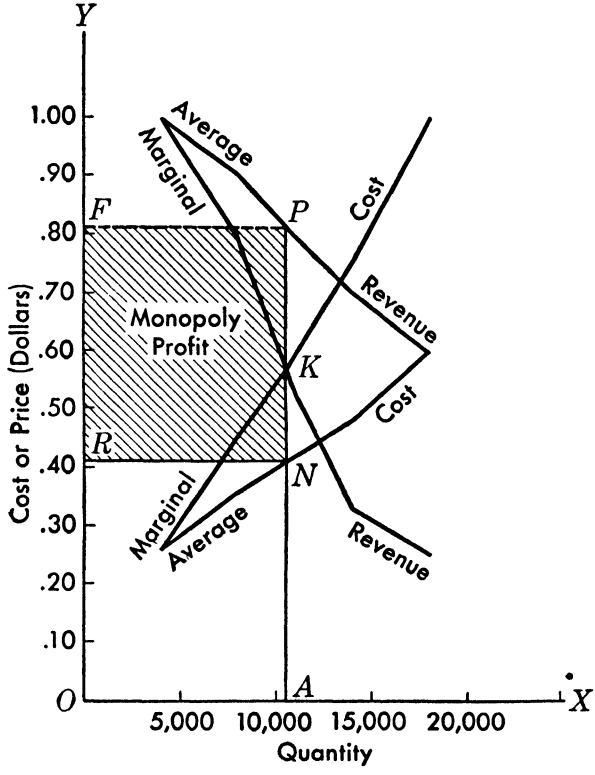


Fig. 4.—Long-run monopoly price under increasing cost.

lowers the per unit price, such an increase also lowers average and marginal unit costs. The problem that confronts him is to determine the combined effect of a decrease in costs, a decrease in price, and an increase in production and sales. Thus, if unit costs decrease rapidly and demand is elastic, in this way causing price to decrease slowly, the monopolist will receive not only a greater profit per unit of output but also a greater total profit because he is selling more units. Moreover, even if his unit costs do not decrease as rapidly as the price, the fact that he will be

able to sell more units may more than offset the lessening of the differential between price and costs.

The following illustrations will aid in clarifying these points. Suppose a monopolist discovers that he can sell 10,000 units at $1 per unit and estimates that he could sell 20,000 units at 60 cents. Let us assume that his average costs are 80 cents when he produces 10,000 units, but that they would be only 42 cents if he expanded his plant and produced

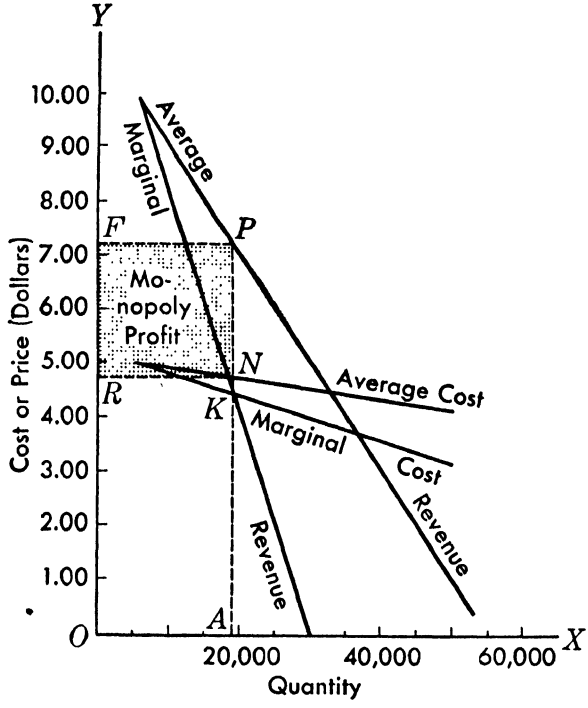


Fig. 5.—Long-run monopoly price under decreasing cost.

20,000 units. This would mean a reduction of 40 cents in price and 38 cents in average cost. Even so, monopoly profits would be increased from $2,000 to $3,600 by increasing production from 10,000 to 20,000 units.

Let us now assume a somewhat different situation. Suppose demand is so elastic that when prices decline from $1 to 60 cents, the quantity sold increases from 10,000 units to 40,000. Assume, further, that costs decrease from 80 cents when 10,000 units are sold to only 50 cents when 40,000 units are sold. We can see that, although price declined 40 cents,

costs declined only 30 cents; yet profits increased from $2,000 to $4,000. Thus, changes in price in relation to cost must also be considered in terms of increases in the number of units sold. Clearly, then, the monopolist operating under decreasing cost conditions is often able to take advantage of both declining costs and increasing volume of sales.

Table 17 and Fig. 5 illustrate the determination of long-run monopoly price under conditions of decreasing cost. The principle involved is the same as before, namely, that production should be expanded until long-run marginal cost is equal to marginal revenue. The reader should now have no difficulty in interpreting either the chart or the table. He should note, however, that under decreasing cost marginal cost is always less than average cost. This means, of course, that when marginal cost and marginal revenue coincide they will both be below average cost. Examination of either the table or the chart will show that under the conditions assumed in this illustration the most profitable price for the monopolist to charge is between $7 and $8.

Table 17.—*Long-run Monopoly Price under Decreasing Cost*

Price	*Units produced and sold annually*	*Total revenue*	*Marginal revenue*	*Total cost*	*Average cost*	*Average marginal cost*	*Total net profit*
$10	5,000	$ 50,000	$ 10	$ 25,000	$5.00	$5.00	$ 25,000
9	10,000	90,000	8	49,000	4.90	4.80	41,000
8	15,000	120,000	6	72,000	4.80	4.60	48,000
7	20,000	140,000	4	94,000	4.70	4.40	46,000
6	25,000	150,000	2	115,000	4.60	4.20	35,000
5	30,000	150,000	0	135,000	4.50	4.00	15,000
4	35,000	140,000	−2	154,000	4.40	3.80	−14,000
3	40,000	120,000	−4	172,000	4.30	3.60	−52,000
2	45,000	90,000	−6	189,000	4.20	3.40	−99,000
1	50,000	50,000	−8	205,000	4.10	3.20	−155,000

Class price. We have been assuming that the monopolist is forced to sell his commodity in the market at one price, but this is often not the case. Once he controls the supply, he may be able to arrange his sales

so that he will secure different prices from people in different income groups. There are many devices used to sell different classes of consumers the same commodity at different prices. The monopolist would make the greatest possible net profit if he were able to sell each additional unit at the maximum price for which it could be sold were no more units available. Although he can never reach this ideal, he can frequently divide consumers into rather large groups and charge different prices to different groups. He can also sell the same product in different market areas at different prices.

The practice of selling the same good or service at different prices is called *price discrimination.* Often this discrimination is possible only if the consumer believes that he is receiving something extra when he buys a good at a higher price. In some cases he does receive something extra, but the extra service or quality that he receives does not justify the difference in the price. Suppose, for example, that a firm has the monopoly of a certain type of safety razor because it holds a patent. It sells its regular model for $5 and a de luxe model for $10. The more expensive model, however, is just like the regular model except for some chromium plating and a special box. Nevertheless, many people who do not have to watch their pennies will gladly pay the extra $5.

Sometimes goods of the same quality can be sold at different prices without making any real distinction. Where this is possible, it usually results from the ignorance of consumers. For example, if a firm had a monopoly of ties, it might sell ties of the same type at different prices in the same market. It might ask $2 for some, $1.50 for others, and $1 for still others. Customers who were used to judging quality on the basis of price would assume that the ties must be different in quality, and hence those accustomed to paying high prices for ties would continue to do so. Meanwhile others would get the same ties at low prices.

However, the monopolist is more likely to practice price discrimination when he sells his product or services in different market areas. Thus he may sell a good in the high-quality stores at one price. At the same time, stores that appeal to a different class of trade are allowed to sell the same good at a lower price. Discrimination in price, therefore, may take place between different markets within the same area. Again, it may occur between different market areas; that is, merchants in one city may

be allowed to sell the commodity at one price, but in another city they may be required to sell it at another price.

Sometimes a producer will have a monopoly in his own country because of a protective tariff. He can therefore charge a high price in the home market, but in the world market he must charge less in order to meet competition. Thus he discriminates in favor of the foreign buyer. The United States has placed high tariffs on a number of commodities; and at times producers have been accused of selling at home at monopoly prices while "dumping" the same commodities abroad at less than average cost.

The possibility of price discrimination in order to increase profits must be carefully considered, for such a practice emphasizes certain principles that are often employed even when monopoly is not complete. We shall therefore begin the analysis with a simple case of price discrimination or class price. We shall assume that a certain monopolist produces but one product, blankets, and that he sells them at different prices, either in the same market or in different markets. We shall also assume that he must decide on his production policy in the intermediate period. He operates a plant with a capacity, under normal conditions, of 200,000 blankets per year. His fixed costs are $200,000 per year, and his variable costs are $1 per blanket produced. His cost situation, for various outputs up to normal capacity, is shown in Table 18.

Table 18

Production	*Fixed cost*	*Variable cost* $1 *per unit*	*Total cost*	*Average cost*
20,000	$200,000	$ 20,000	$220,000	$11.00
50,000	200,000	50,000	250,000	5.00
100,000	200,000	100,000	300,000	3.00
150,000	200,000	150,000	350,000	2.33
200,000	200,000	200,000	400,000	2.00

Table 19 supplements Table 18. It shows the demand for blankets, the total revenue, total cost, and total profit at each price. It also shows the

additional blankets that the monopolist can sell whenever he reduces the price.

If the monopolist were able to charge only one price, he would restrict his output to 100,000 blankets for the year and sell them at $4.50 each. His total profit would then be at the maximum of $150,000. Suppose, however, to make the illustration as simple as possible, that he sells various units of his commodity in different markets. If he charges $15 per blanket, he can sell in only one market, and his demand is limited

Table 19

Price	*Quantity demanded*	*Total revenue*	*Total cost*	*Total profit*	*Additional blankets demanded*
$15.00	20,000	$300,000	$220,000	$ 80,000	
7.00	50,000	350,000	250,000	100,000	30,000
4.50	100,000	450,000	300,000	150,000	50,000
3.00	150,000	450,000	350,000	100,000	50,000
2.00	200,000	400,000	400,000	0	50,000

to 20,000 blankets. By reducing the price to $7 in another market, he can sell an additional 30,000 blankets. Likewise each reduction in price can be made to conform to the ability to pay of individuals in different markets until all the blankets are sold.

Our example, however, indicates that in order to operate his plant at capacity he was forced to sell some of his stock at the average cost of production when his plant was operating at capacity. The question then is, Has he increased his total net profits by this procedure? We must observe that when he produces at normal capacity his total costs are $400,000. But what about his total revenue if he sells at class prices? We can find this in Table 20 by adding up the revenues received from the sale of blankets at each of the successive prices indicated.

Since the total cost at capacity production is only $400,000 while the total revenue is $985,000, the total net profit is $585,000. This is $435,000 more than he was able to make by selling at one price. However, not all monopolists are in so fortunate a position. Sometimes the monopolist

will not be able to sell all that he can produce at prices that will cover his average unit cost. Let us see how he would fare under a different demand and price situation. This is represented in Table 21.

Table 20

Quantities sold at each successive price	*Price*	*Total revenue from each quantity*
20,000	$15.00	$300,000
30,000	7.00	210,000
50,000	4.50	225,000
50,000	3.00	150,000
50,000	2.00	100,000
200,000		$985,000

Under the conditions of demand indicated in Table 21 $4.50 is still the price that will give the highest total net profit. However, if production and sales are increased beyond 100,000 blankets, the price drops very rapidly. If we turn back to Table 18, we note that if the monopolist produces to capacity his average unit cost will be $2. It is clear, therefore, that, if he wishes to produce at capacity under the conditions

Table 21

Price	*Quantity Demanded*
$15.00	20,000
7.00	50,000
4.50	100,000
2.00	150,000
1.10	200,000

indicated in Table 21, he will be obliged to sell some of his product at prices below average cost. How will such sales affect his profits?

In Table 22 profits are calculated on the basis of the relation of price to the *average* cost of $2 per unit that would prevail at capacity production.

Table 22

Quantity sold at each successive price	*Average cost at capacity*	*Price*	*Total revenue from each quantity*	*Total cost of each quantity*	*Total net profit or loss on each quantity*
20,000	$2	$15.00	$300,000	$ 40,000	$ 260,000
30,000	2	7.00	210,000	60,000	150,000
50,000	2	4.50	225,000	100,000	125,000
50,000	2	3.00	150,000	100,000	50,000
50,000	2	1.10	55,000	100,000	−45,000
Total profit on all sales					$ 540,000

The table shows that the monopolist makes a total net profit of $540,000. However, if profit is calculated on the basis of average cost at capacity output, he suffers a loss of $45,000 on the last 50,000 blankets. Since this is the case, why not stop production at some other point? Suppose the monopolist produced only 20,000 blankets and sold them at $15 each. His total income is $300,000; but since he is not producing at capacity, his actual total costs are the $220,000 shown in Table 19 rather than the $40,000 shown in Table 22. Consequently, his total net profit is only $80,000. If he sells 50,000 blankets, his total income is $510,000 and his total costs are $250,000; thus his total net profits are only $260,000. Again, if he continues selling at discriminatory prices and sells another 50,000 at $4.50, he increases total income to $735,000, but his costs are increased to $300,000 and his profits are still only $435,000. Again, suppose that he sells another 50,000 blankets at $3 per blanket. He will increase total revenue to $885,000; and since his costs will be $350,000, he will make a net profit of $535,000. This profit, however, still falls $5,000 short of the amount that he will make if he sells another 50,000 blankets, even though he sells the last 50,000 below the average cost of production.

Two points should be noted in this analysis of the action of the monopolist: (1) By producing to capacity he lowers the average cost per unit for all units and thus increases the apparent profit per unit on all except the last 50,000 units. (2) Although he has an apparent loss of

$45,000 on the last 50,000 blankets, he is able to sell these above his variable costs, which in this case are also his marginal costs. Since his marginal costs are only $1 and the price is $1.10, he actually gains 10 cents per unit by selling the last 50,000 blankets. As long, therefore, as the monopolist can sell additional units above variable costs, he can increase his profits by expanding production. This assumes, of course, that he can maintain his class prices on the other units.

The principle that the monopolist employs may be stated as follows: Where possible classify customers and sell to each class at the highest price they will pay. As long as excess capacity remains, keep reducing prices to reach new groups, provided that the lowest price received at least covers variable costs of production.

Summary

A producer has a monopoly if he is the sole seller of a commodity or service. In the utility industries monopoly is the rule, but prices are under government control. In other fields complete monopoly is the exception, and even where it exists it is likely to be difficult to maintain over a long period of time. Moreover, even a complete monopoly is subject to the competition of substitutes and, if the substitutes are good, such competition may be very effective.

In spite of these limitations on monopoly there are two reasons that make it important to understand how monopoly prices are determined: (1) Minor monopolies like those arising from patents or copyrights are fairly common. (2) If one understands how prices are determined under both perfect competition and complete monopoly, he can better understand the great number of cases falling in between these two extremes.

A study of price determination under complete monopoly indicates that, just as in the case of competition, cost of production plays a less important part in the short run than it does in the intermediate period or the long run. The monopolist can exert control over price because he can control the quantity of goods placed on the market. In the short run he will be concerned principally with the relation of price to total revenue. However, for both the intermediate period and the long run he must consider the effect of price changes on both revenue and costs. In either period his total net profits will be greatest if he expands production until marginal cost and marginal revenue are equal.

The monopolist, however, seeks to increase total net profits by another means, namely, by classifying his customers and charging discriminatory prices as between classes. In this way he is often able to increase his profits by increasing revenue without increasing costs. Moreover, if by charging class

prices he can utilize his plant more fully, he may be able to reduce average unit costs. To secure capacity production the monopolist may even sell some of his product at prices that are below average unit costs. This increases his total profit so long as these prices are not below variable unit costs.

It has already been noted that markets which approach perfect competition or complete monopoly are the exception rather than the rule. The following chapter will deal with some of the many cases falling in between these two extremes.

Study Questions

1. What is meant by a fixed supply?
2. Explain why the average revenue curve is the same as the demand curve.
3. Will a higher price always bring the monopolist a higher profit?
4. Under what conditions will monopoly be found when the supply is reproducible?
5. How is marginal revenue determined? Contrast marginal unit revenue and marginal unit cost.
6. Show why profits of the monopolist are always greatest when the marginal unit revenue just equals the marginal unit cost.
7. Show the relationship of the intersection of the marginal revenue and marginal cost curves to the average cost curve under conditions of constant cost, increasing cost, and decreasing cost.
8. What is meant by class price?
9. Name some of the methods used by the monopolist to secure a class price.
10. If a monopolist charges a class price, under what conditions is it possible to secure a higher net profit even when some units of the commodity are sold below total unit costs?

Exercises

1. Mr. A owns a rare painting. He has received the following offers: $20,000, $40,000, $60,000, $80,000. What will the price be? Can this situation be represented by demand and supply schedules?

2. What price would the monopolist charge if the following conditions prevailed?

Unit price	*Units sold*	*Revenue*			*Costs*				*Profit or loss*
		Total, absolute	*Marginal*		*Total*		*Marginal*		
			Absolute	*Unit*	*Absolute*	*Unit*	*Absolute*	*Unit*	
$0.80	1,000	$ 800.00	.....	.....	$ 700.00				
0.75	1,200	900.00	$100.00	$0.50	720.00				
0.70	1,400	980.00	80.00	0.40	750.00				
0.65	1,650	1,072.50	92.50	0.37	842.50				
0.60	1,900	1,140.00	67.50	0.27	942.50				
0.55	2,200	1,210.00	70.00	0.233	1,077.50				
0.50	2,500	1,250.00	40.00	0.133	1,232.50				

Draw a chart showing monopoly profits.

3. A monopolist has a plant capable of producing 100,000 units. His total unit cost, when operating at capacity, is 50 cents per unit; variable unit costs are 20 cents. The monopolist discovers that he can sell 70,000 units in the United States at $1 per unit. Because a tariff has been imposed by the government, he is protected in the home market. He decides, therefore, to sell in a foreign market, but, because of competition, he is forced to sell the 30,000 units at 30 cents per unit, a price that is below total cost when operating at capacity.

However, if he produces only 70,000 units, his total unit cost is 70 cents per unit. Should the monopolist sell in the foreign market, even though he sells below total unit costs, or should he be content with his home market; that is, in which case will he make the highest profit?

a. If the monopolist can charge but one price for total supply produced, how many units will be produced and what will be his profit?

b. Suppose that he is able to charge a class price and can separate customers so that he can charge each group a price as indicated above. How much profit will he make?

Unit price	*Units sold*	*Revenue*			*Costs*					*Profit or loss*
		Total, absolute	*Marginal*		*Absolute*			*Marginal*		
			Absolute	*Unit*	*Fixed*	*Var.*	*Tot. un.*	*Abso-lute*	*Unit*	
$1.00	20,000	$20,000	.....	.....	20,000	4,000	24			
0.80	40,000	32,000	$ 12,000	$0.60	20,000	8,000	28			
0.70	70,000	49,000	17,000	0.567	20,000	14,000	34			
0.60	110,000	66,000	17,000	0.425	20,000	22,000	42			
0.50	160,000	80,000	14,000	0.28	20,000	32,000	52			
0.40	220,000	88,000	8,000	0.133	20,000	44,000	64			
0.25	250,000	62,500	−25,500	−0.85	20,000	50,000	70			

4. A monopolist is operating under the following market conditions: He operates a plant with capacity to produce 250,000 units. His total fixed costs are $20,000; his *variable costs are* 20 *cents per unit* produced.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 15.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 14.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 12.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 13.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 15.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 20.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chaps. 10 and 11.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chap. 18.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 17.

17. Imperfect Competition

Both perfect competition and pure monopoly are rare in real markets. The usual situation is what we can call *imperfect competition,* or that type of market condition where there is more than one seller but where each firm producing and selling a product exercises some control over price, even if very slight. In other words, competition is not perfect, but neither does pure monopoly exist. The competitive situation is somewhere between these two extremes. We shall concentrate our attention on two types of markets which fall under the general classification of imperfect competition. In one of these, the number of sellers is small and therefore the control over price may be great. This type of market we shall refer to as *partial monopoly.* In the other type there may be a relatively large number of sellers. However, competition is by no means perfect in spite of the large number of sellers, because each firm makes a product that differs slightly from that made by any other. There is an important degree of competition between the sellers, nevertheless, because the product of each firm serving the market fills the same general need of the consumer. This second market situation we shall refer to as *monopolistic competition.*

We shall begin our investigation of imperfect competition with an analysis of partial monopoly.

Partial Monopoly

Partial monopoly prevails in a market when a single commodity is being produced by only a few firms. These sellers compete if they have no agreements with each other. But each seller must consider the reaction of all the others in planning his price and production policies, because he realizes that each of his few competitors can affect prices and production in the market.

Duopoly. The closest approximation to monopoly is found when there are but two sellers in the same market, both offering the same commodity at a single price. Both firms are often about equal in size; hence they must respect each other's power to compete. Although pure duopoly, like pure monopoly, rarely exists in the sense that there are only two firms in the market, each producing a single commodity and selling it at a single price, sometimes two sellers so dominate the field that duopoly is said to exist. For example, the TNEC, created by Congress to investigate the concentration of economic power, states,[1] "In the electric lamp business, General Electric shares a duopoly of metal bases with Westinghouse and a duopoly of large glass bulbs, glass tubing, and rods with the Corning Glass Works." The report mentions other fields in which two firms dominated, even though in each of these fields there were other sellers. These other sellers, however, were so small in comparison with the dominant firms that they could not influence price very much.

In order to study the action of duopolists, let us take a simple example. Suppose there are but two clothing merchants, A and B, in a certain market, and both sell the same brand of clothes. Under these conditions it is clear that neither seller can charge a price that varies much from that charged by his competitor. If A, for example, decides to lower his price, those who formerly purchased suits from B will now buy from A. The loss of trade by B will force him to cut prices if he is to recoup his trade. A knows that this is the only logical action by B. He knows that B may retaliate by slashing prices below that now charged by A. Such a procedure, however, would result in a still further lowering of price by A to keep his trade. A continuous cutting of price in retaliation can result only in destructive competition costly to both sellers.

Since both sellers, A and B, realize the implications of lowering prices to secure business from a competitor of about equal size, each will hesitate to start price competition. Each seller, moreover, is interested in making a profit. The problem confronting both A and B is how to secure the highest total profit in the long run without a price war. The answer lies in establishing a price that will maximize profits for both sellers. Each will seek to maintain this price when it is found. They can then divide profits between them. Thus, the price under duopoly will tend to approximate the price that will be charged by a monopolist, for the monopolist will charge a price that maximizes total net profits.

[1] "Competition and Monopoly in American Industry," Monograph 21, p. 199.

We should note one point, however. Although each is willing to charge a price that will maximize profits of both, the duopolists cannot without an agreement approximate this price as closely as a monopolist who can survey his market and regulate production of the entire output. Therefore the price under duopoly will tend to approximate monopoly price, but it may be slightly higher or lower, since control of output is not absolute. It is probable, moreover, that price under duopoly will be a little lower than that charged by a monopolist, because each duopolist fears competition from the other if he raises price in an effort to maximize profit.

Strictly speaking, duopolists are two firms producing the same product. However, it might be to their advantage to produce slightly different products that fill the same need of the consumer. In other words, it would be possible, and indeed probable, that product differentiation would occur.

Oligopoly. A more common condition found in the market is that of several competing producers with capacities of about the same size. Their number, however, is so small that no one of them dares act independently of the others in his pricing and production policies. Thus we talk about the "Big Three" in the automobile industry. Again, we find that in the chemical field three firms, du Pont, Allied Chemical and Dye Corporation, and Union Carbide and Carbon Corporation, dominated the market in certain lines before the war, although there are several other corporations only slightly smaller than these three leaders in size and influence. Hundreds of small companies are also in this field and some of them made great progress during the Second World War in expanding products. Rayon is also said to be dominated by three firms, du Pont, The Celanese Corporation of America, and American Viscose, but in this industry also there are several other large firms. In the tobacco industry the dominance of a few firms is likewise in evidence. We can see, therefore, that many products are produced under conditions that approximate oligopoly.

Let us take a simple example to illustrate long-run price determination under oligopoly. Instead of two clothing firms dominating a certain market, assume that four have a leading position; that is, each firm can influence the market price by changing the price at which it sells its commodity. We again shall assume that all firms sell only one brand of

suits and that all must sell at a single price because they are competing. Under these conditions, what is the position of the oligopolist?

Although the oligopolist realizes that any change in his pricing policy will affect all his competitors and may bring retaliatory actions from them, it is more difficult to determine how each competitor will react to a price change. The oligopolist knows that if he lowers his price all others will be forced to follow, for otherwise they will lose some of their customers. But what about increasing price? No one seller can be sure that if he increases price other sellers will follow; if they do not, he loses customers to them. Furthermore, the greater the number of firms in the market, the less any one seller will know about the others. Although each seller can influence the market and each desires to enhance profits, the fact that no seller can be sure of the action of the others means that price is not likely to be increased to a point where total net profits will be maximized. Hence price under oligopoly, although inflexible, is not likely to be so high as under duopoly, just as price under duopoly may not be so high as that charged by a monopolist.

We should note that one characteristic is common in all three cases: prices tend to be inflexible. Under monopoly a seller can set the most profitable price and maintain it; under duopoly and oligopoly a seller is reluctant to change price because he fears the reaction of his competitors. We may, therefore, conclude that under partial monopoly price is not so high as under monopoly, but that price is likely to be inflexible because it does not rapidly respond to changes in demand or cost.

We need not confine our discussion of oligopoly to a situation where identical products are sold. Product differentiation can be present here also. When a duopolist or an oligopolist is able to differentiate his product even slightly, he may be able to secure a larger share of the market without causing destructive price competition.

Where duopoly exists, the advantages of monopoly may be so great that the two firms will enter into an agreement and become a monopoly. In the United States, however, there are legal obstacles to doing this openly. Where there are more than two firms in the market, as under oligopoly, the effect of competition is stronger. Price will therefore approximate that which would prevail were competition more nearly pure, and there is less danger of collusion on the part of all of the sellers. Later in this chapter we shall apply these generalizations to real market conditions.

Thus far in the discussion of duopoly and oligopoly, we have emphasized the influence of the seller in the determination of prices. Large buyers, however, also have their influence and can force sellers to quote prices that are satisfactory to them. Thus the great mail-order houses wielded great power in certain markets, such as refrigerators and tires, just before the Second World War. Moreover, even though these commodities were produced by oligopolists, in the sense that only a few important sellers were in the market, the power of the large wholesale buyer was so great that the partial monopoly power of the producers could not always be asserted.

Partial monopoly in practice. The analysis has pointed out how those who are duopolists or oligopolists may be expected to act in applying their pricing policies. It has also pointed out that, because of fear of retaliation, price charged by such sellers may be inflexible; that is, it does not change readily with shifts in demand. Often, in fact, under conditions of partial monopoly, competition is limited to other considerations than price; and producers may be willing to let sales decline in order to maintain a certain price. We can see that those who have a partial monopoly will use every device at their command to increase their control of production. Thus they often use product differentiation as well as production control. Partial monopolists, however, generally manufacture more than one commodity; hence, price policies and costs must be considered in relation not only to a single commodity but also to all commodities sold. In this section we shall consider actual cases to determine some of the methods that have been used by partial monopolists.

Examples of duopoly. The TNEC, which investigated the concentration of economic power, reported that in 1936 duopoly was found in the following commodities: domestic telegraph service, international communications, bananas, plate glass, electric lamps, electric accounting machines, air brakes, oxyacetylene, and sulfur. The committee also asserted that many public utilities operate as duopolies, and that often duopoly prevails in local markets, but it pointed out that data were not available to determine to just what extent duopoly actually existed.[2]

The committee also pointed out that often firms that have a partial monopoly in one phase of their operations may have a complete monopoly in others. Thus their policies result, not from partial monopoly alone, but also from other avenues of price control. In considering the action of the

[2] *Ibid.*, pp. 98-112.

duopolist, therefore, we must remember that he operates in a complicated economic system. Often he produces many commodities, some of which are produced under actual monopoly conditions, others under actual competition. Then too, he can practice product differentiation and thus affect prices. Often, also, he is a buyer as well as a seller, and consequently he can impose his power upon other sellers as well as upon buyers. Finally, in some cases he may enter into a definite agreement with his competitor.

In order to see how the duopolist operates, let us take the case of bananas, as reported by the committee. This committee stated that, although bananas are grown in all moist tropical countries, more than nine-tenths of those produced for the export market come from Central and South America and the West Indies. The United States is one of the greatest consumers of bananas. In 1936 this country imported over 99 per cent of its supply from Central America. Two concerns handle the bulk of this trade: The United Fruit Company, which sells about 60 per cent, and the Standard Fruit and Steamship Company, which sells 30 per cent.

United is, therefore, the giant in the industry. In 1936 it owned or leased 3½ million acres of land. It operated hundreds of ships. It owned all but one of the banana railroad lines in Central America. It owned the only telegraph company in several Central American countries. It operated docks, stores, hospitals, hotels, and thus dominated the economic life of some of these states.

Even so, it produced only about one-half the bananas it sold, and it purchased the other half. This company would not contract with a producer unless he would agree to sell his entire supply to it. Since the company owned and operated railroads and docks, it could charge recalcitrant producers high rates for the use of these services or it could exclude them entirely.

Its only competitor, the Standard, according to the report, "prefers to work with United rather than oppose it." Standard did not compete at all in several Central American states. To reciprocate, United did not compete with Standard in several others. Although in some states both purchased bananas, each concern often purchased all of the product within an area. In such areas neither company competed with the other.

We may expect that, with two firms in the market and with one occupying a dominant position, the leader can exercise power like that

exercised by a monopolist. We have seen that the two firms in our illustration did not actually compete, for Standard apparently allowed United to set the price and then acted accordingly. Under these conditions, United was free to use a policy similar to that employed by monopolists; namely, it limited the supply of fruit. Sometimes bananas were cut and allowed to rot in the field; they were purchased and dumped along the railroad tracks; they were thrown into the ocean.[3] By these methods the retail price of bananas was kept at a level profitable to the sellers.

A study of the other commodities mentioned by TNEC indicates that the duopolists in almost every case used tactics similar to those employed by United. Since, however, most of the other duopolists were manufacturers, it was not necessary for them to destroy the good but simply not to produce it. The evidence showed that often duopolists tended to form agreements and to use other devices to maintain satisfactory prices. Thus their action was similar to that of a monopolist.

Examples of oligopoly. Some examples of oligopoly were mentioned above, namely, automobiles, chemicals, and rayons. Many other commodities, however, are also produced under conditions of oligopoly. The TNEC listed fifty-four industries in which the four largest firms produced more than two-thirds, by value, of the total supply. Included in the list were gum, ammunition, cigarettes, linoleum, rubber tires and inner tubes, sewing machines, soap, and matches.

Some firms, however, produce more than one commodity. It is, therefore, not at all uncommon to find the same firm dominating several fields. Often the same stockholders control several firms. Thus the leading firm in the cement industry in 1936 was the Universal-Atlas Cement Company, which is a subsidiary of the United States Steel Corporation, a leading company in the production of steel. In the case of steel, even with several firms in the field and with only three or four dominant, the largest firm is still so large that it overshadows its nearest rival. In 1936, the United States Steel Corporation was several times as large as its nearest rival, Bethlehem; and the latter in turn was nearly twice as large as the third largest firm, Republic Steel Corporation. This latter corporation, however, was so large that its assets exceeded the aggregate investment of most of the remaining major firms put together.[4]

[3] *Ibid.*, p. 102.

[4] *Ibid.*, p. 119.

Where such firms exist and each firm produces more than one commodity, the power of each firm over price will vary greatly among the various commodities it produces. Yet, in spite of the complexity of the economic system in which these firms function, we can trace certain trends. In fact, sellers tend to act in most cases as our theory of oligopoly would lead us to expect. That is why it is important to understand basic principles of this market situation.

For instance, under oligopoly if one firm dominates the industry, its competitors are unlikely to engage in open competition with this dominant firm. Rather, the other firms tend to follow its leadership. In other words, when the dominant firm raises or lowers price, all firms in the industry also change price in the same direction. This is called *price leadership.* Such price leadership was found in the steel industry. It was also present in agricultural implements when the industry was dominated by the International Harvester Company, which manufactured 41 per cent of the total farm machinery. In oil, the successors of Standard Oil invariably took the lead in announcing price changes; in copper and lead, the American Smelting and Refining Company assumed this role; and so on throughout the list of the commodities that were produced by a few leading firms, among which one firm, because of its size, dominated the industry. This action is what we would expect in theory, and this is what was found in practice.

A second expected result was found in most of these industries. Large competitors reached agreements to control the amount placed on the market and thus avoid spoiling the market by forcing a reduction in price. The committee showed in its monograph that agreements were reached in iron ore, gasoline, chemicals, typewriters, and other commodities.

A third situation, which would be expected under oligopoly, was found among heavy goods industries, such as steel, cement, cast-iron soil pipe, and similar heavy products, where transportation charges are a large part of total cost. Firms in these industries by agreement used the basing-point system. They quoted prices for delivery, which included transportation costs from a certain base point. For many years most steel companies used the "Pittsburgh plus" basing system whereby every firm, no matter where it was located, quoted price plus freight charges from Pittsburgh. The effect was to make the cost of steel to the buyer the same whether he bought from a producer in his own area or from a distant producer. In 1924 the government issued orders that this system was to be dis-

continued. Since that time, instead of having one basing point, several have been used.

Under the new basing-point system as it was operating in 1936, all those who sold steel charged the same price as United States Steel. They then calculated the freight charges from various basing points and selected the one that gave the lowest price. However, since all sellers used the same formula, every seller of any steel product quoted the same price to the buyer for delivery at his door.[5] This system was varied somewhat, of course, in other industries, but the result was the same; namely, identical prices were quoted by all sellers, irrespective of their location.

Another practice used when a few firms dominate the industry is *market sharing*. This may be considered as a fourth result of oligopoly. Under market sharing, the dominant firms divide up the market in order to avoid competition. Sometimes such sharing is by formal agreements; at other times this practice is carried on as a trade policy. In the monograph mentioned above, the authors point out that this practice has in the past characterized investment banking, anthracite coal, meat, tobacco, and other commodity fields.[6]

The monograph contains the following clear statement of the attitude of sellers who seek to avoid price competition: [7]

> It should be observed that the general practice among intelligent competitors respecting one another's position need not be a matter of "tacit understanding." In the case of Swift and Company it is an individual common-sense policy, arrived at independently, not to invite retaliation and trade wars by using aggressive tactics. Swift has deliberately tried to avoid cutthroat competition wherever it was legally possible to do so.

In order to avoid competition, various other devices are also used. Common ownership and control of enterprises in the same industry are not uncommon. Interlocking directorates, common interest groups, and trade associations may be other means by which partial monopoly is preserved.

We must not assume that those able to exert partial monopoly power always do so. The examples cited show what certain partial monopolists have done in the past. Legislation, court decisions, and changes in trade

[5] *Ibid.*, p. 150.
[6] *Ibid.*, p. 188.
[7] *Ibid.*, p. 185.

practices have modified some activities of duopolists and oligopolists since the TNEC made its study.

In the industrial history of the nation other examples may be cited, however, showing how dominant firms in an industry have sought to destroy one another to secure monopoly power. Instances can be found in which sellers, plagued with idle capacity, were willing to take the risk of incurring destructive competition by lowering prices and thus increasing the amount they could sell. In many cases these sellers have engaged in price wars only to discover that such tactics were not so profitable as cooperation. It is important to know that, just as the advantages of monopoly have led to many agreements, so the pressures of competition have often led to price wars.

The rubber tire industry has been a good example of highly concentrated control with vigorous price competition rather than agreement. Five firms, Goodyear, Firestone, U. S. Rubber, Goodrich, and General, dominate the field. Four of these firms are located in the same city. The product, however, is largely one for which the demand is indirect; the sale of tires depends upon the number of cars in operation and the number of miles the cars are driven. During the years, moreover, better and better tires have been made, and with the improvement of roads the number of tires sold per car has been reduced. Prices of tires have fallen. in spite of improvements. Tires that sold for $23 in 1920 sold for $4 in 1932, and their average life had been greatly increased. The low price in 1932, however, was due in large measure to the depression.

Of further significance in this industry was the fact that productive capacity was increasing faster than demand; hence, there was idle capacity. The industry suffered from excess capacity, rapidly falling prices, and advanced technology. These factors caused many of the smaller firms either to go into bankruptcy or to enter into mergers with other firms in the hope of cutting overhead. Notwithstanding the gradual elimination of small firms, in spite of mergers, and in spite of the fact that there were only five dominant firms in the industry, four of which were in the same city, there was little evidence of collusion or agreement to fix prices or divide the market.

Several factors account for the continuation of competition in an industry that has every reason to become monopolistic in order to save itself from destructive competition. First was the policy of Mr. Firestone, who directed the affairs of the second largest tire-producing company.

He believed that the best way to increase volume was to reduce prices. Since his plant was operated very efficiently, his costs were low; consequently, he continued to reduce prices, and his action forced other manufacturers to follow suit.

However, this factor was not the only one that affected competition in the tire industry. Though the industry was under the domination of a few firms, it was faced with another problem. Automobile manufacturers were among the largest buyers of new tires. Their business was, therefore, keenly solicited by the sellers. This fact made it possible for automobile manufacturers to play one seller of tires against another and thus force price concessions. Replacement of tires was also largely in control of a few large buyers, such as the great mail-order houses, oil companies, and auto-supply chain stores. These large buyers also forced tire companies to make concessions in prices. In the main such concessions resulted in the manufacture of special brands for the seller. Thus Goodyear made All-state tires for Sears, Roebuck and Company and sold them to Sears at a price from 29 to 40 per cent below the wholesale price of Goodyear's comparable brand, All-weather tires. Sears in turn sold its All-state tires from 20 to 25 per cent below the prices charged by Goodyear dealers for a comparable tire. Under somewhat similar conditions United States Rubber Company was making Riverside tires for Montgomery Ward, while Goodrich was making Atlas tires for Standard Oil. Tire companies, moreover, also sold cheaper brands of their own in competition with the brands that they were making for large buyers. Thus we can see that competition was very great between these manufacturers. After 1933, however, such competition was greatly modified, and price cutting, dealer pirating, and the like seem to have been largely abolished.

Price wars and severe competition have taken place among other industries in which a few firms dominate the market. The refrigerator field has had its price wars; food processors dominating their fields have kept prices flexible because of their active competition. On the other hand, many examples can be given to show how concentration of production in a few firms has kept prices rigid and profits high. Drugs, medicines, soaps, cosmetics, and toilet preparations have generally been sold under a system of inflexible prices and high profits.

Monopolistic competition. A second type of imperfect competition, which we have called *monopolistic competition,* has easily defined characteristics, the most significant being product differentiation. As we

have mentioned, there may be a large number of sellers, but each offers a slightly different product. We shall now examine the ways by which the products of competing firms can be differentiated even though these products serve the same general need of the consumer.

How product differentiation can be achieved. Product differentiation can be achieved in a number of ways. In the first place, products may be differentiated because they have different specifications or are made from different ingredients. For example, there are a relatively large number of radio manufacturers who make table models. These radios differ considerably in specifications. Another example is found in the large number of tooth pastes available on the American market. Most of them differ in some degree because of the various ingredients used in their manufacture.

When product differentiation results from differences in specifications or ingredients, the differentiation is real. However, the consumer sometimes may be persuaded that one product is superior to others even when no real differences exist or when the differences are very slight. The persuasion may result from various types of selling efforts.

The selling efforts of the cigarette companies provide an example of differentiation through persuasion. Cigarettes are made from tobacco and cigarette paper, and all are classified as cigarettes. However, various sellers seek to differentiate their product from that of others. One seller maintains that only finer tobaccos are used; another, that his product contains less nicotine; a third, that the cigarette he produces burns longer and is less irritating to the throat; a fourth, that something new has been added. Each producer also competes with the others in making supposedly scientific tests and in securing testimonials from noted persons to make the consumer believe that his product is different from that sold by competitors.

Thus advertising and various other types of selling efforts can be used to achieve product differentiation. There are many other ways by which the goal can be reached, however. Within any market some sellers have unusually desirable locations. Other long-established firms have built up a reservoir of good will that causes customers to look with special favor on their products. Still other enterprises provide such excellent and courteous service that their products are given preference even when the goods are identical in physical characteristics. The use of brands and trade-marks is a means that contributes to establishing and maintaining

product differentiation. A brand name or trade-mark helps the consumer identify the product that he has been convinced is superior. The value of a trade-mark or brand name usually is built up by long and expensive advertising.

Why product differentiation is sought. In the degree to which product differentiation and preference can be attained, a firm can increase the extent of the monopoly it enjoys. This in itself eliminates some of the hazards of losing customers through price cutting by competitors. Moreover, higher prices usually can be obtained when product superiority can be established. Thus product differentiation is highly desirable in a market where there is a relatively large number of sellers.

The student need only to look around to observe that monopolistic competition is the most common type of market situation. There are many suppliers of most products that we buy. Each seller tries to make us believe that his product is different and superior.

When differentiation is established, a firm becomes a partial monopoly in dealing with its most loyal customers. As with any monopoly, its sales curve slopes downward to the right. However, its position is not nearly so secure as would be the case if it had a complete monopoly. Its power over price is limited by the fact that somewhat similar products are offered by competitors. An increase in price will drive some customers to competitors if the competitors hold prices unchanged. On the other hand, by lowering prices customers can be won from other concerns; unless, of course, the others meet the price cuts.

Summary

Since both perfect competition and pure monopoly are rare in real life, we need to understand price determination under what we termed *imperfect competition.* By imperfect competition we mean that type of situation where characteristics of both monopoly and competition are found. There is more than one seller, but the individual sellers nevertheless wield considerable influence in regard to price. Most real markets are examples of imperfect competition.

We recognized *partial monopoly* and *monopolistic competition* as two types of imperfect competition. We further subdivided partial monopoly into duopoly, where there were only two sellers, and oligopoly, where relatively few sellers were present. The chief characteristic of monopolistic competition

is *product differentiation;* that is, the product of each of a relatively large number of sellers is different in a slight degree from that of each of his competitors. Differences in product can occur, of course, in partial monopoly, but are not the outstanding characteristic of this type of market situation. Rather, the small number of sellers is the dominant feature.

Enough illustrations have been cited to show the problems that arise when sellers enter a market under the complex conditions of production and of sales that prevail in the actual world of business. We have seen that, even where producers are selling one commodity only, each is confronted by competitive and monopolistic forces which vitally affect him. Sometimes the seller stands in a favorable position to agree with his competitors to keep prices high, only to discover that, because of vast buyer power, he must make concessions. Sometimes sellers fear government regulation. Then, too, low-cost producers are not always willing to share the market with high-cost producers. To some degree, however, each seller through the use of brands can differentiate his product and thereby secure and retain a market. Generally such a market can be retained only at a very great advertising cost.

Although a producer may be tempted to enter into agreements to prevent price cutting and many other competitive practices, the pull of competition may be so strong that these agreements cannot be sustained; yet they are attempted again and again, only to be interrupted by severe price wars and destructive competition. Notwithstanding these difficulties, we can predict long-run tendencies or trends. For example, our study has shown that duopolists and oligopolists tend to keep prices rigid and to restrict production in order to secure high profits. This long-run trend may be interrupted time and again; yet it persists. Hence, we find our theory substantiated in actual practice; and we may be confident that where duopoly or oligopoly exists the trend will be toward rigid prices, collusion, and agreements that may eventually lead to monopoly. On the other hand, in markets dominated by competition, we discover the wide use of brands and product differentiation. These devices prevent the operation of free competition. However, notwithstanding these trends, the pull of competition is still a vital force in the market.

Study Questions

1. What two phases of imperfect competition are studied in this chapter? Distinguish between them.
2. What is the difference between monopoly, oligopoly, and duopoly?
3. From the point of view of theory, what would you expect to happen to the price-cost relationship in a market in which the number of firms

producing a particular product increased from two to five? From five to ten? From ten to one hundred?

4. What objectives may sellers have in mind when product differentiation is attempted?

5. Name several methods used to secure product differentiation.

6. Under what form of market (monopoly, perfect competition, oligopoly, etc.) would you expect the greatest price flexibility, other factors being equal? Why?

7. What factors does the oligopolist weigh in considering a change in pricing policy? Do these same factors confront the monopolist?

Exercises

1. Our study has indicated that where partial monopoly exists, prices tend to be much more inflexible than where competition dominates.

In 1938 the automobile industry published the results of a study entitled, "The Dynamics of Automobile Demand," [8] in which the writer stated:

> There is a certain school of economic thought that characterizes the automobile industry as an "oligopoly." According to this school, each producer controls such a large proportion of the total market that he must reckon with the effect of his own actions upon the actions of his competitors and consequently is restrained from cutting prices because he knows that a price reduction would promptly be met by his competitors. Therefore, subscribers to this doctrine maintain that prices fail to fall in oligopolistic industries in response to sudden reductions in the intensity of demand, whereas in industries in which there is little concentration, prices respond very quickly to changes in intensity of demand. This analysis shows that the above line of reasoning, at least as applied to the automobile industry, is specious. Concentration has nothing to do with it but costs have everything to do with it, and the nature of costs is entirely independent of concentration of production in a few hands.

a. To what extent does the discussion in the text agree with the statements made in this article relative to the action of oligopolists in controlling prices? To what extent does it disagree?

b. Does the discussion in the text emphasize the significance of costs as a price-determining factor in oligopoly? In what way?

c. The writer also presented figures showing that prices had changed greatly between 1920 and 1939. Thus car A was priced at $3,170 in 1920 and a

[8] "The Dynamics of Automobile Demand," Detroit, Mich., 1938, pp. 133-134.

comparable car was priced at $685 in 1939. Prices of car B were $2,435 and $648, respectively; car C $3,285 and $810; car D $2,780 and $955.

Does the fact that over the years prices of a commodity are radically reduced invalidate the theory of price under oligopoly?

2. Some economists maintain that producers of commodities that are fundamentally the same seek to make them appear different by use of the following means: (1) use of clever slogans, (2) differences in conditions of sale, (3) differences in sales policies, (4) differences in advertising.

a. Study certain products that are fundamentally the same and determine how many of these devices are being used. For example, such products as cigarettes, soap, and shaving equipment are typical commodities in which product differentiation is practiced.

b. Would you classify automobiles in this group of commodities? Why or why not?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 16.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Part I, Chaps. 18-28.

Chamberlain, E. H.: "Theory of Monopolistic Competition," Harvard University Press, Cambridge, Chaps. 1-5.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chaps. 12 and 13.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 13.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 12.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 21.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 18.

Robinson, Joan: "The Economics of Imperfect Competition," Macmillan & Co., Ltd., London, 1933, Chaps. 1-4.

Stigler, G. J.: "The Theory of Competitive Price," The Macmillan Company, New York, 1942, Chaps. 5-9.

18. Interdependent Prices

In studying the forces that determine the prices of individual commodities we have thus far assumed that the price of any one commodity is determined independently of the prices of other commodities. Strictly speaking, this assumption is not tenable. The prices of all commodities, both services and material goods, constitute a *price system;* and the price of any one commodity is related in various ways to the prices of all the others.

In the first place, all prices are interrelated because all economic goods compete directly or indirectly for the consumer's dollar. As a result, a change in the price of any one good, and the total amount of money spent on it by consumers, will tend to affect in some degree the demand for most other goods, and hence their prices. Let us suppose, for example, that the price of automobiles should rise 50 per cent because of reduced supplies of labor and materials. Unless the demand for automobiles has unit elasticity, this would cause people to spend on an average either more or less of their income on cars. If demand is elastic, they would spend less; if it is inelastic, they would spend more. The national income remaining unchanged, if people spent more money for automobiles they would have less to spend on other goods; and therefore the demand for other goods would tend to fall off, and their prices would tend to decline. On the other hand, if people spent less money on automobiles, they could have more to spend on other goods, demand for these goods would tend to increase, and their prices would tend to rise.

A second way in which the prices of economic goods are related to one another is through competition for the factors of production. Suppose, for example, that there is an increase in the demand for some one product, say airplanes. Airplane prices will rise and airplane production will be more profitable; and as a result the producers will attempt to expand output. To do this they will have to employ more of the factors of production. They will, for one thing, hire more labor. This increased demand

for labor for producing airplanes will tend to raise the general level of wages; and as a result it will tend to raise the costs and prices of all other products.

It is clear, therefore, that theoretically a change in the price of one commodity will tend to affect the prices of all other commodities. In most cases, however, the effect on other prices is so small that we make a very slight error if we assume, as we have in earlier chapters, that the price of a given commodity can change independently of the prices of other commodities. There are, however, some important cases in which the prices of two or more commodities are so closely tied together that any considerable change in the price of one is likely to have a marked effect on the prices of the others. The remainder of this chapter will be devoted to a study of some of these special cases of interrelated prices.

Prices under joint demand. When from necessity or custom two or more products are used together, and when a change in the demand for one is almost sure to be accompanied by a similar change in the demand for the others, the products involved are said to be subject to *joint demand.* Under actual market conditions there are many cases of this kind. For example, if the demand for bread increases, the demand for butter will increase; if there is a greater demand for automobiles, then there will be a greater demand for gasoline and oil; and if there is an increase in the demand for houses, this will create increased demands for bricks, stone, glass, wood, nails, and other materials used in building.

When two commodities are subject to conditions of joint demand, a change in demand will cause the prices of both to move in the same direction. However, if the price of one of the commodities changes as a result of a change in supply, the price of the other will tend to move in the opposite direction. Let us suppose, for example, that the price of bread should rise to 50 cents a loaf because of a reduced supply resulting from an extreme shortage of wheat. The high price of bread would cause a drop in consumption; and since butter is principally used on bread, the reduced consumption of bread would cause a drop in the demand for butter, and hence a decline in the price of butter.

Where several commodities are subject to conditions of joint demand, the demand for any one of them considered separately is likely to be relatively inelastic. An extreme example is the demand for nails. It has

already been noted that a demand for houses creates a joint demand for nails and various other building materials. Let us suppose that the demand for houses is very elastic; so much so that a rise of 10 per cent in costs will cause a decline of 50 per cent in the volume of construction, while a decline of 10 per cent in costs will cause a doubling of the volume of construction. In spite of this elasticity of the demand for houses, the demand for nails alone will be highly inelastic. Even if the price of nails should increase ten times, it is unlikely that sales would fall off very much. The reason is that nails must be used if houses are built; and though the increase in the price of nails will increase the cost of houses, the effect will be so slight as to have no appreciable effect on the number of houses demanded.

It has already been noted that, when two commodities are subject to joint demand in a certain use, a change in this joint demand will tend to cause their prices to move in the same direction. However, if one of these commodities has only this single use while the other has many uses, the price of the first commodity will be affected much more than the price of the second. Such a situation is often found when the demand for a finished good creates a joint demand for the factors used in producing it.

To illustrate such a case as simply as possible, let us assume that two factors of production, and two only, are jointly demanded for the manufacture of a finished product. We will suppose that only wood and labor are required for the making of chairs, and that a demand for chairs creates a joint demand for both of these factors. At a certain time when the market is in equilibrium, the wood used for each chair costs $1 and the labor 50 cents, while the chair itself sells for $1.50, a price just sufficient to cover the cost. At this price 10,000 chairs are demanded per month.

Now let us suppose that there is a substantial increase in the demand for chairs so that the price rises to $2, at which price the number demanded is 15,000 per month. This increase in the demand for chairs will increase the demand for both wood and labor and will tend to increase the prices of both. Assume, however, that wood is used in many industries and that the demand for it in chair making is a very unimportant part of the total market demand. In this situation the increased demand from the chair industry will have little effect on the price of wood. Assume, on the other hand, that the type of labor used in making

chairs is highly skilled and specialized and is not employed in any other industry. The increased demand for chairs will mean a great increase in the demand for this specialized type of labor and will tend to cause a sharp rise in its price. As a result, when equilibrium is again reestablished so that the price of chairs is once more equal to their cost of production, most of the increase of 50 cents in the price of chairs will go to labor in the form of higher wages, and very little of it will be represented by an increase in the price of wood.

This example illustrates how, under conditions of joint demand, a change in the demand for the finished product may have different effects on the prices of different factors of production. We see that, where joint demand exists, variations in the prices that producers must pay for the factors of production will be greater for factors with few uses than for those which have many uses. However, the differences between these two situations may be less in the long run than in the short run. If because of scarcity there is a sharp rise in the price of some factor, in the long run this is likely to bring about an increase in the amount supplied. Given time, for example, new men can be trained in the art of making chairs. Thus extreme changes in the prices of the factors used in production may prove to be only temporary.

Prices of substitutes. When two or more economic goods may be used for the same purposes, they are said to be *substitutes*. Substitute commodities are very common, not only among consumers' goods, but also among materials used in production. Sometimes two substitute goods are equally desirable for the purpose that they serve, and in this case the tendency is always to use the cheaper one. More frequently, perhaps, one of the substitute goods is better than the other, and in this case the less desirable substitute will be used only if the saving in cost is appreciable.

When two goods can be substituted for one another, a change in their relative desirability tends to cause their prices to move in *opposite* directions. For example, butter and oleomargarine are substitute commodities but most people prefer butter to oleo. However, if the public preference for butter should decline, the demand for butter and its price would tend to fall while the demand for oleo and its price would tend to rise.

The result is different when the price of one of two substitute goods changes because of a change in its supply. In this case the price of the

other good will tend to move in the *same* direction. Suppose, for example, that an epidemic among dairy cows should cause a decline in the supply of butter and a sharp rise in its price. It is clear that this would cause an increase in the demand for oleo, and that the price of oleo would also rise.

When a commodity has one or more good substitutes that sell within its own price range, the demand for it is likely to be more elastic than would otherwise be the case. If the price of such a commodity rises, the amount demanded will decline rapidly because people will turn to the substitutes. On the other hand, if its price falls, the amount demanded will increase rapidly because people will use less of the substitutes and more of the commodity that has declined in price.

Because substitutes for a commodity make its demand more elastic, they tend to reduce fluctuations in its price. Consider once more the case of butter and oleo. A decline in the supply of butter will cause its price to rise. However, the fact that people can turn to oleo as the price of butter rises causes the rise in price to be considerably less than it would be if no oleo were available as a substitute.

Often the manufacturers of a finished product can choose from among a number of substitute raw materials. The maker of chairs, for example, can employ any one of several kinds of wood, such as oak, maple, or mahogany. He can even, if he chooses, make chairs of steel or aluminum, or of some metal alloy. To be sure, the various substitutes available are not equally desirable from the viewpoint of all purchasers of chairs. Nevertheless, a rise in the price of chairs made of any one material will cause many consumers to turn to chairs made of substitute materials. As a result, the prices of chairs are likely to fluctuate considerably less than would be the case if substitute materials were not available.

Partial, complete, and potential substitution. When each of two substitute commodities is in active use, substitution may be said to be *partial.* It sometimes happens, however, that eventually one of the two commodities entirely replaces the other. In this case substitution is said to be *complete.* Complete substitution may result because one good develops an overwhelming superiority for the use in which both have been employed; or it may occur because one good becomes so very much cheaper that no one is willing to pay the price of the other. It is clear that, when a substitute commodity completely replaces its rival, the displaced commodity ceases to exert any active influence on the price of

the commodity still in use. The automobile and the horse and buggy are an example of almost complete substitution. At one time both were in wide use, and the price of one undoubtedly influenced the price of the other. Now, however, the superiority of the automobile is overwhelming, and it has almost completely replaced the horse and buggy. As a result, there is no longer any significant price relationship between the two. Today if the price of automobiles should go extremely high, many people would go without them but very few would turn to the horse and buggy as a substitute.

In many cases where a commodity has no active substitutes there may be materials that could be employed in its place if conditions changed. Where such a situation exists, substitution is said to be *potential.* The existence of potential substitutes for a commodity is sometimes very significant. For example, the known petroleum reserves of the United States will not last very many years. However, satisfactory motor fuels can be made from various other materials at prices not very greatly above the present price paid for gasoline. The existence of these potential substitutes for petroleum gives the public protection against a possible future shortage of gasoline and the excessive prices that this would bring.

Prices of joint-cost products. It often happens that the production of one good results in the production of other goods. For example, a farmer who raises cotton also secures cottonseed. When sheep are raised, the product is not only meat but also wool and hides; and when copper is produced, a certain amount of silver is extracted from the ore as a by-product. Whenever it is impossible to produce one commodity without at the same time producing others, we have a true case of *joint costs.* When two commodities are produced under conditions of joint cost, it is obvious that there is no way of separating the cost of one from the cost of the other.

Let us take as a simple illustration the production of cotton and cottonseed. Suppose that there is a demand for each of these commodities and that each can be sold at a price. How much should the farmer receive for cotton fiber if he is to cover its cost? And how much should he receive for cottonseed? Since the cost of producing the fiber cannot be separated from the cost of producing the seed, it is obvious that these questions cannot be answered separately. The only thing that can be said is this: If it is to be worth while for the farmer to go on raising cotton, the price

of the cotton fiber and the cottonseed together must be at least high enough to cover all the economic costs of producing both.

There was a time when cottonseed had no value and was thrown away as waste. In this situation the entire cost of producing cotton had to be carried by the price of the fiber. Finally, however, cottonseed was put to various uses. The seed became a source of vegetable oil, and cottonseed cake was used for feeding purposes. Since cottonseed had then become useful, for the first time it commanded a price in the market. As a result, cotton raising became more profitable than before. This, however, tended to result in greater production of cotton and lower prices for both the fiber and the seed. The tendency was for the prices of both to drop until the price of the two together was just equal to the joint cost of producing both. To the farmer it makes no difference how much of this joint cost is covered by the fiber and how much by the seed. If he can cover his total economic costs, his most difficult problem is solved.

To see the nature of the problem confronted by the cotton farmer let us suppose that it costs on an average 10 cents to raise a pound of cotton when cotton fiber alone is produced, and that the price of a pound of cotton is also 10 cents. Assume that later a use is found for cottonseed, and that for every pound of cotton produced there is also a pound of cottonseed. If the cottonseed sells for 5 cents a pound, the farmer will receive 15 cents for a pound of cotton fiber plus a pound of seed. Since his total costs are only 10 cents, he will be induced to increase production. The output of a single farmer will not, of course, affect the price. Other farmers, however, will also have an inducement to expand output, and the result will be that the prices of both cotton and cottonseed will decline. If costs do not increase, the expansion of production will continue until the combined price of a pound of cotton and a pound of cottonseed has dropped to their joint cost of 10 cents. The individual farmer can influence neither prices nor the general level of costs. He can, however, strive to operate his own farm with more than average efficiency in the hope of bringing his own costs somewhat below the general level.

When two commodities are produced under conditions of joint cost, their *relative* prices will be determined entirely by demand. Let us assume as before that it costs 10 cents to raise a pound of cotton and a pound of cottonseed, and let us further assume that each product sells for

5 cents a pound and thus bears half of the joint cost. Sometime later there is an increase in the demand for cotton fiber but no increase in the demand for cottonseed. Suppose that the increased demand for cotton fiber causes its price to rise to 7 cents. This higher price will induce an increase in the production of both the fiber and the seed. However, the increase in the output of cottonseed will cause its price to decline, let us say to 3 cents a pound. Since the price of cottonseed has declined and the price of the fiber has risen, a greater part of the total cost of production is now covered by the price of the fiber. Whereas previously the fiber and the seed each covered one-half the joint cost, now the price of cotton fiber covers 7 cents of the 10 cents' cost of producing a pound of each, while the cottonseed covers only 3 cents of this cost. However, so long as the total cost is covered, it is possible to continue to produce both products at the established rate of output.

Under perfect competition, it should be noted, the combined prices of joint-cost products would always tend to equal the total cost of production, but the relative prices of such products would be determined solely by demand.

Before leaving the discussion of joint costs, we should note that there are few cases in which *all* the costs of producing two or more products are joint. Beef and hides, for example, are subject to joint costs only up to the point where the cattle are slaughtered and skinned. Beyond that point special costs must be incurred before either product is ready for the final consumer. Unless the price is high enough to meet these special costs, the product will be worthless for it will not pay to market it. In the case of hides the principal special costs that must be met are those involved in curing, tanning, and shipping.

Prices of inseparable-cost goods. Inseparable costs must be carefully distinguished from joint costs. As has already been noted, a true case of joint costs exists only when it is impossible to produce a certain commodity without producing one or more other commodities. There are many cases, however, in which commodities that could be produced quite independently of one another are in fact produced by the same enterprise or even in the same plant. Where two or more commodities of this type are made in the same productive unit, there is no way of allocating the "overhead" or fixed costs except on a purely arbitrary basis. Such costs are therefore said to be *inseparable.*

To illustrate some of the problems that arise in connection with inseparable costs, let us suppose that a certain manufacturer is producing several products and that substantial parts of his costs are inseparable. Since there is no satisfactory way of allocating these inseparable costs, how can he tell whether any particular product is being manufactured at a profit or at a loss? The answer to this question is that a product is always profitable, in the sense of increasing his net income or reducing his net loss, if the price of the product more than covers the separable costs incurred in producing it. These separable or direct costs are the only costs that would be saved if the product in question were not manufactured. If the price of the product more than covers these separable costs, then producing it contributes something toward the overhead or fixed costs that go on anyway.

Of course in the long run a manufacturer must cover inseparable as well as separable costs. He should not produce any commodity, even at a profit, if he can replace it with another that will yield a still greater profit. Assuming, however, that a more profitable alternative is not available, it pays to produce any commodity that can be sold at a price that will more than cover the separable or direct costs of production.

Charging what the traffic will bear. A very large proportion of the costs of a railroad are of the overhead or inseparable type because of the huge investment in plant and the heavy cost of maintaining the right of way. As a result, a railroad faces certain peculiar problems in settling its rates. On an average its rates must be high enough to cover all costs. However, if the same flat rate per ton-mile were charged for all goods regardless of their type or value, some cheap and heavy products, like gravel, might not be shipped at all. In such cases it will sometimes pay a railroad to set a rate lower than its average total costs per ton-mile if such a rate will result in bringing it more traffic. Of course if it carries some kinds of freight at rates below average total costs, it must carry other kinds at rates above average total costs. It therefore places relatively high rates on light and valuable goods which will stand these charges. Such a method of setting freight rates is called "charging what the traffic will bear."

The following illustration will help the reader to see why it may be profitable for a railroad to carry some kinds of freight at less than average total unit cost. Suppose that the average cost to a railroad of carrying all freight is 2 cents per ton-mile. Gravel, however, will not be

shipped if the rate is over 1½ cents. Suppose that the separable or direct costs incurred in carrying gravel are only 1 cent a ton-mile. In this situation if the railroad refused to carry gravel for 1½ cents, it would save a cost of only 1 cent and sacrifice a revenue of 1½ cents for every ton-mile of traffic lost. For every ton-mile of traffic gained, it adds ½ cent to its net income. Even the shippers of goods that pay freight rates above the average are likely to gain, for if the railroad sacrificed the net income to be obtained from gravel it would have to charge even higher rates on other types of goods.

"Charging what the traffic will bear" has a sinister sound to many people, because at one time some railroads used their monopoly power to charge excessive and oppressive rates. Today, however, charging what the traffic will bear means only charging those rates which will lead to the fullest and most profitable utilization of plant. Because such a policy reduces average costs, it is likely in the long run to benefit equally both the railroads and those who use their services.

Summary

The prices of all goods are interrelated because all goods compete for the consumer's dollar or for the factors of production. However, in the case of most commodities the error is small if we assume that the price of a given commodity can change independently of the prices of other commodities. There are, however, certain special cases in which the prices of two or more commodities are rather closely tied together. In this chapter four of these special cases were analyzed: joint-demand goods, substitute goods, joint-cost goods, and inseparable-cost goods.

Joint-demand goods may be defined as goods that, from custom or necessity, are used together so that a change in the demand for one is accompanied by a change in the same direction in the demand for the others. Substitute goods are goods that may be used for the same purposes. Joint-cost goods are goods that cannot be produced separately. If one is produced, the others are produced in the same process. Inseparable-cost goods are goods which could be produced separately but which are frequently produced in the same productive unit. As a result, a large part of the costs of producing them consists of overhead or inseparable costs which cannot be allocated to any particular product except on a purely arbitrary basis. The fact that railroads are subject to large overhead costs accounts for their practice of setting rates on the basis of "what the traffic will bear."

Study Questions

1. In what respects are all prices related?
2. Why does an increase or decrease in the price of one commodity affect prices of other goods?
3. Under what conditions do we have a joint demand?
4. When several products are in joint demand, what effect will an increase in the supply of one have upon the prices of the others? Why?
5. When a number of products are jointly demanded, the demand for any one of them is likely to be rather inelastic. Why?
6. When is one good a substitute for another?
7. How does the existence of substitute goods affect prices? Explain carefully.
8. Under what conditions will prices of substitute goods move in the *same* direction? in the *opposite* direction?
9. What is the difference between partial, potential, and complete substitution?
10. What is the difference between joint demand and joint cost?
11. Show how the price of one commodity produced under conditions of joint cost is affected by an increase in the demand for another commodity. Explain all possible situations.
12. Explain the difference between joint costs and inseparable costs of goods.
13. Under what conditions should prices be based on the principle of charging what the traffic will bear?

Exercises

1. The transportation industry includes airplanes, motor vehicles, trains, boats, and privately operated automobiles. Show how the fact that all these means of transport exist affects the amount of traffic hauled by each; the price each can charge; and the profits each can make. Is one of these industries a complete substitute for the other? a partial substitute? a potential substitute?

2. Copper and silver are often produced from the same ore. During the last several years the government has pursued a policy of keeping the price of silver high.

a. Suppose silver and copper are produced in equal amounts from the same ore. What effect will this policy have on the quantity of copper produced?

b. Suppose for every ounce of silver produced, 10 pounds of copper are produced. Suppose that under these conditions silver sells for 70 cents per ounce while copper sells for 20 cents per pound. Will your answer still be the same?

c. What effect will a decrease in the demand for silver have upon production under the above conditions? What effect will a decrease in the demand for copper have? Explain your answers carefully.

3. Suppose that the demand for pens and ink represents a joint demand. Assume that pens are produced under conditions of decreasing cost in the long run, and that ink is produced in the long run under conditions of increasing costs. Suppose that in the long run the demand for pens increases greatly. What effect will this increase in demand have upon the price of pens in the long run? upon the price of ink?

4. In some states a very heavy tax is placed on oleomargarine. What effect will such a tax have on the price of butter? Explain.

5. Wood, steel, and aluminum are used to make many articles that are used for the same purpose. Show how an increase in the cost of production of any one of these goods will affect the demand for the others.

6. Does the principle of substitution modify pricing theory? Explain.

7. How would you determine how to charge what the traffic will bear? Could you make such a charge if you were operating under conditions of competition? Explain. Under monopoly? Do you think that department stores in pricing goods ever use this principle?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 14, pp. 285-290.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 16.

DEIBLER, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 9.

FAIRCHILD, F. R., E. S. FURNISS, and N. S. BUCK: "Economics," The Macmillan Company, New York, 1940, Chap. 14.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 11.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 11.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chap. 15; Vol. II, Chap. 63.

19. Economics of International Trade

Price determination up to this point has been discussed in terms of one uniform market such as might be found within the boundaries of one country. Now we are prepared to go beyond national borders and to study the processes by which goods are exchanged between nations. The question of international trade is destined to become of increasing importance to Americans as more and more millions of our citizens owe their employment to the willingness and ability of people in other lands to buy our goods and services.

Since our government began to encourage world commerce with the reciprocal trade treaties of the 1930's, it has been pointed out frequently that we must export much more than we have in the past if we are to have full employment at home. There is more to a flourishing trade than the volume of our exports, however. In order to have a vigorous export trade, our citizens must learn that we also must import more than we have previously. This principle, that we must import if we would export, seems obvious to students of economics, but it is not always accepted. Some groups of our citizens are willing enough to sell abroad, but they urge high tariffs to prevent foreigners from competing in our home markets. Since America is the greatest market for goods as well as the leading producer, the whole world, as well as Americans themselves, will benefit as our citizens realize that world trade is not a one-way street.

Differences between domestic and international trade. In dealing with international trade, we encounter problems that are not met when goods are sold in the domestic market. In the first place, each nation has its own system of money. The monetary unit of the United States is the dollar. In Mexico it is the peso; in England it is the pound. The expert

in international trade must understand the relative values of the dollar and the French franc, or of the Norwegian krone and the Japanese yen. In short, he must have a thorough knowledge of rates of exchange, which we may define as the number of units of one country's money that will purchase one unit of another nation's money. For example, the currency exchange rates shown in Table 23 were declared official by the International Monetary Fund at the time that organization announced its first quotations in late December, 1946.

The problem of tariffs also arises in connection with international trade. A tariff is a form of tax which is levied by a government on goods imported into a country. It may be a flat fee per unit of goods or it may be ad valorem, in which case the tariff is based on the value of the goods. Some nations levy tariffs solely to collect revenue. Other nations purposely levy restrictive tariffs, or tariffs whose main objective is to prevent or impede the importation of goods. These usually are referred to as protective tariffs. A tariff established to raise revenue naturally is not high enough to prevent the importation of goods; otherwise goods would not be imported and no revenue would be collected.

Sometimes protective tariffs are established at rates that are designed to equalize the costs and prices of goods imported and those produced at home. When this occurs, domestic consumers are prevented from obtaining goods abroad at lower prices than home producers can supply them.

Another problem also must be considered in connection with international trade. During the 1930's, many nations hindered the importation of goods by setting quotas on what could be brought in from particular countries or by limiting the amount of funds that their own citizens could use to purchase foreign goods. The objective of such quotas was to balance the volume of imports and exports, or to limit the amount spent for ordinary imports so that military supplies could be purchased abroad. One of the goals of the International Monetary Fund is to eliminate such impediments to world trade.

As a fourth problem of international trade, the ordinary negotiations of buying and selling are complicated by differences in language and business customs in the different nations. Such barriers are not consequential, however. In fact, a flourishing world trade encourages a better understanding between nations and a greater appreciation by the people of all countries of the culture and institutions of other peoples.

Table 23

Country, currency	*Par values in terms of U. S. dollars*	
	Currency units per U. S. dollar	*U. S. cents per currency unit*
Belgium, franc	43.828	2.282
Bolivia, boliviano	42.000	2.381
Canada, dollar	1.000	100.000
Chile, peso	31.000	3.226
Colombia, peso	1.740	57.143
Costa Rica, colon	5.615	17.809
Cuba, peso	1.000	100.000
Czechoslovakia, koruna	50.000	2.000
Denmark, krone	4.799	20.838
Ecuador, sucre	13.500	7.407
Egypt, pound	0.242	413.300
El Salvador, colon	2.500	40.000
Ethiopia, dollar	2.484	40.250
France, franc	119.107	0.840
Guatemala, quetzal	1.000	100.000
Honduras, lempira	2.000	50.000
Iceland, krona	6.488	15.411
India, rupee	3.308	30.225
Iran, rial	32.250	3.101
Iraq, dinar	0.248	403.000
Luxemburg, franc	43.828	2.282
Mexico, peso	4.855	20.597
Netherlands, guilder	2.652	37.695
Nicaragua, cordoba	5.000	20.000
Norway, krone	4.963	20.150
Panama, balboa	1.000	100.000
Paraguay, guarani	3.090	32.363
Peru, sol	6.500	15.385
Philippine Commonwealth, peso	2.000	50.000
Union of South Africa, pound	0.248	403.000
United Kingdom, pound	0.248	403.000
United States, dollar	1.000	100.000

Advantages of world trade. Before entering into a discussion of the processes by which goods and services are exchanged between nations, it would be well to list the advantages of world trade. In the first place, world trade permits regional specialization based on climate and natural resources. Bananas can be raised more advantageously in Central America than in Canada, which in turn can produce cheaper paper. It is therefore sensible for Canada to exchange paper for bananas rather than to raise the bananas at home. Some countries, again, can export and others must import metals such as nickel and copper, because the important natural deposits are concentrated in a few areas. Second, trade between nations permits national skills and aptitudes, developed over long periods, to be utilized to their fullest capacity. Americans make the best automobiles and machinery in the world. The Swiss have developed great skill in watchmaking and in processing dairy products. In the third place, trade permits occupational specialization. Nations with large populations but with small areas of agricultural land can specialize in manufacturing and then exchange their fabricated goods for the agricultural products of more sparsely settled areas.

Closely related to occupational and regional specialization is the principle of diminishing returns. If nations are willing to engage in trade with each other, each nation can gear its productive mechanism to its most efficient rate of output. So long as additional supplies of a commodity can be imported at reasonable cost, a nation need not push the production of any commodity into the area where costs rise sharply because of diminishing returns. Thus, world trade permits people of all nations to enjoy abundance of goods and services at the lowest possible cost.

Attitude of nations toward world trade. The policies of various nations toward foreign trade have been expressed in well-developed doctrines which have been given much study by economic historians. At various periods of history, certain doctrines have gained ascendancy. During the nineteenth century the doctrine of *free trade* was accepted by many of the leading commercial nations. Under a policy of free trade either tariffs were not collected when goods crossed national boundaries, or they were collected solely for revenue. As already mentioned, tariffs for revenue were not high enough to discourage trade, because the policy of advocates of free trade was to encourage international commerce and competition between nations.

Before Adam Smith and some other economists of his period took up

the cudgels for free trade, the prevailing doctrine was known as *mercantilism.* The proponents of this doctrine advocated a "favorable balance of trade." This meant to them that their nation should export more than it imported, with the difference being paid for in gold. It was never fully explained how this objective could be accomplished if every nation tried to sell more abroad than it purchased. The basis of the doctrine was that wealth consisted of precious metals rather than of useful goods and services. One of the most ardent advocates of mercantilism was Jean Baptiste Colbert, the great French financier of the seventeenth century. Colbert, as a part of his national policy, sought to make France self-sufficient as far as imports were concerned. At the same time he tried to add to the national hoard of gold and silver by selling as much abroad as he could. The folly of his policy was brought home then, as it has been so many times since, when Colbert got his nation into a costly war with the Dutch and made enemies of the other trading nations.

In between these two extremes have been various other attitudes toward trade. Many of the principles of mercantilism were revived during the period between the two world wars when so many countries sought to achieve economic self-sufficiency through high tariffs which would build up home industries and eliminate the necessity of imports of certain commodities. The policy of economic nationalism, which called for building a strong, self-sufficient national economy that was not dependent on trade for essential goods, sometimes carried the exclusion of imports far beyond what the mercantilists of the sixteenth and seventeenth centuries would have advocated. During the 1930's some European nations barred certain kinds of imports entirely by fiat. Some set quotas on imports that could be brought in from particular countries. Such quotas usually were related to exports to these other nations. Others accomplished the same objective by prohibitive tariffs. Germany helped build her synthetic rubber industry, for example, by establishing such a high tariff on natural rubber that the cost of the synthetic product could be recovered when sold in competition with the natural product within the country. Some of the arguments for and against high tariffs will be presented later in this chapter.

The stake of the United States in world trade. Considerable concern has been expressed over the apparent chronic export surplus of the United States. American tariffs have been high, natural resources of the United States are great, and America is the principal source of supply for a world that needs vast quantities of machinery and other capital

goods. The strong position of the United States as the workshop of the world was enhanced by the destruction of Japan and Germany, our great commercial rivals before the war. Only Great Britain is a rival now and her capacity to produce for export is limited.

The huge demand for United States goods was demonstrated in the immediate postwar period when new peacetime records for exports of American merchandise were set in 1946 and again in 1947. Our imports also established new records but were far below the volume of exports. In order to finance the excess of exports the United States granted loans or made outright gifts amounting to billions of dollars. The sale of gold and American securities held by foreigners also provided dollars for purchase of American goods. The various emergency aid programs also swelled United States exports in the postwar period since this country was almost the only source of supplies and equipment.

A plan was proposed by Secretary of State Marshall under which funds were to be provided by the United States government for aiding European nations over a period of years. This was said to be necessary because it became clear that these nations would not be able to export enough immediately to pay for the imports needed to reconstruct their economies. Such a method of providing assistance would mean that the export surplus of American merchandise could continue after other resources for making purchases in this country had been exhausted.

Pessimists in regard to the outlook for our international trade recall that the total value of our exports, including reexports, dropped from an average monthly total of $436,750,000 in 1929 to only $134,251,000 in 1932. This decline occurred, of course, during the great world-wide depression that began in late 1929, but our pessimistic friends point out that after we stopped lending, as we did so freely in the 1920's, we did not regain until the Second World War the level of export business that we had when we were making loans with which others could buy our goods. It also is pointed out by those who scan trade statistics that, previous to 1946, our all-time record in foreign trade for a full peacetime year was reached in 1920 when the sum of our exports and imports totaled about 13 billion dollars, or only a little less than the volume of 1946. In 1920 we were making heavy loans to foreign nations just as we were in 1946. Those who take the pessimistic view concerning our prospects in world trade point to the heavy defaults of the First World War debtors to us. They also note that 45.8 per cent of foreign dollar bonds

recently were in default. We are told that we can export almost any quantity of goods we wish if we are willing to give them away or accept promises to pay that eventually will prove worthless.

The real argument of the pessimists is that our international trade is fundamentally out of balance and will be so even after the adjustments from war to peace have been completed. Such an opinion hardly seems justified. In the 16 years before Munich when normal peacetime trade relationships prevailed, the United States exported 58.1 billion dollars' worth of merchandise and imported 50.4 billion dollars' worth. A number of net payments by Americans to foreigners made up the difference. During these years Americans spent 3.7 billion dollars more on foreign travel than foreigners spent in this country for travel. We had a net deficit of 1.2 billion dollars for shipping and freight services. Our net deficit on miscellaneous minor transactions was 1.6 billion dollars, and personal remittances of American immigrants to relatives abroad amounted to 3.6 billion dollars. Thus these deficit items, most of which were "invisible," actually exceeded the entire merchandise surplus during the 16-year period. Later in this chapter "invisible" and "visible" exports and imports will be discussed.

The scope of our foreign trade and its importance to every American community can be seen by the volume of exports of a number of different items taken at random from the whole range of goods sent abroad. The following figures show the average monthly volume of exports of a group of items in the years 1936 to 1938, when normal peacetime relationships prevailed to a large degree:

Passenger automobiles	$ 9,388,000
Trucks and busses	6,308,000
Electric refrigerators	1,604,000
Radio apparatus	2,319,000
Office appliances	2,684,000
Raw cotton	26,059,000
Meat	2,182,000
Wheat	3,296,000
Lumber	3,832,000
Paint and varnishes	1,611,000
Copper and copper manufactures	6,419,000
Metalworking machinery and machine tools	5,891,000
Iron and steel products	12,397,000
Paper and paper manufactures	2,207,000
Photographic and projection goods	1,757,000

Agencies designed to foster world trade. Most businessmen, economists, and statesmen have realized that the world cannot achieve prosperity unless world trade is revivified in the postwar period. During the depression years and the war years, the nations of the world left the gold standard and it appears unlikely to be revived. The gold standard gave a measure of stability to trade. It provided a means of establishing exchange rates between nations that adhered to it. When the gold standard was functioning smoothly, the businessman who sold goods abroad could be fairly sure of knowing that he would receive in payment a sum equivalent in purchasing power to the sum agreed upon when the sale was made. This was not always true after the nations left the gold standard. Rates of exchange fluctuated violently on occasion, and the hazards of trade were increased. As a hypothetical example of losses that can occur from shifts in exchange rates, a Detroit automobile company might be owed 1,000 guilders by a Dutch importer. After receiving payment, the motor company might wish to exchange guilders for dollars. We shall assume that the 1,000 guilders would be worth about $376 at one rate of exchange. If the Dutch government should reduce the value of its currency to half, the 1,000 guilders would be worth only $188 in U. S. funds. In fairness to the Dutch, it should be said that they would be unlikely to slash the value of their currency so drastically. They are well known for their soundness in finance.

The nations of the world have set up two agencies designed to foster stability in international economic relations: the International Monetary Fund and the International Bank for Reconstruction and Development. Both came into existence on Dec. 27, 1945, when agreements governing them were signed by sponsoring nations in Washington.

The purposes of the International Monetary Fund bear directly on the problems of trade outlined above. The fund is intended to lessen the hazards of foreign trade and to prevent countries from winning temporary advantage by unexpectedly changing the values quoted for their paper money. The 44 nations that were members of the fund when it began operation in 1947 are pledged not to alter the value of their currencies without forewarning the fund. Each member contributes to the fund so that a pool of capital is established which will facilitate the transfer of payments between nations when the usual commercial channels cannot provide the means. The fund will extend short-term credits to countries to help them over temporary difficulties in meeting

payments that they must make to other nations. The fund provides for readjustment of exchange rates when such adjustments are warranted.

A member may obtain foreign currencies from the fund in exchange for its own currency to assist in meeting deficits in its international payments for current transactions. Limits are set on a member's purchases of foreign currencies from the fund, and these limits may not be exceeded except by special permission. Under usual conditions, a member may purchase foreign currencies with its own currency until the fund's holdings of its currency equal 200 per cent of its quota, that is, 200 per cent of the amount that it originally subscribed to the fund. After the fund's holdings of its currency reach 75 per cent of the quota, they may not be increased in any 12-month period by more than 25 per cent of the quota.

Actually, most dealings in foreign exchange will pass through the usual banking channels just as now. Resources of the fund merely supplement the ordinary foreign exchange resources of the member nations and constitute a reserve that can be called upon to meet temporary difficulties.

The total amount subscribed to the fund by the member nations shortly before it began operations was $7,397,500,000. The quotas of the principal nations and the percentage of the total of each quota are shown in Table 24.

Table 24

Country	*Amount in millions of dollars*	*Per cent of total*
United States	$2,750	37.2
United Kingdom	1,300	17.6
China	550	7.4
France	450	6.1
India	400	5.4
Canada	300	4.1
Netherlands	275	3.7
Belgium	225	3.0
Latin America, total	469.5	6.3
Other countries	678	9.2
Total	$7,397.5	100.0

The Bank for Reconstruction and Development, which had 38 members in early 1946, was organized to aid reconstruction in war-devastated nations and to promote economic development in backward countries. It will provide funds for realization of these objectives by direct loans from its own capital, by direct loans from funds raised by selling its own bonds, and by guaranteeing bond issues sold by approved foreign borrowers in the private capital markets. It is hoped that, when backward nations are developed through aid supplied by the bank, they will be able to take their places in world trade. Further, the speeding of reconstruction work through aid of the bank will permit war-ravaged nations to return to their places in trade. Subscriptions to the bank by leading nations are shown in Table 25.

Table 25

Country	*In millions of dollars*	*Per cent of total*
United States	$3,175	41.4
United Kingdom	1,300	17.0
China	600	7.8
France	450	5.9
India	400	5.2
Canada	325	4.2
Netherlands	275	3.6
Belgium	225	2.9
Latin America, total	288	3.8
Other countries	632	8.2
Total	$7,670	100.0

Basis for trade between nations. In discussing the nature of international trade, we have already observed that some areas can produce goods at lower cost than can others. Oranges will not grow outside of hot houses in northern climates, nor are gold mines or nickel deposits found in every country. Under these conditions, trade would normally take place when a nation cannot produce a good that it needs or cannot produce it as cheaply as other nations can. The only problem here is a

possible attempt by the country containing the product desired to exploit the country that does not have it.

During the First World War, owners of nitrate deposits in South America greatly increased the price that they charged. Since the United States needed nitrates for explosives, Americans were forced to pay these high prices. The only recourse possessed by the United States was to discover a method of producing synthetic nitrates in order to break the nitrate monopoly.

Other examples can be cited. The British and Dutch, who controlled rubber, sought in the 1920's to exploit the United States, which had no rubber. Today synthetic rubber promises to prevent any further attempts to exploit this country so far as this commodity is concerned. Japan once controlled the entire output of camphor and used its control to exploit the world until a substitute was discovered. Most of the world's diamonds are produced in South Africa, and the policy is to withhold them from the market when the price starts downward.

Thus the problem becomes one of preventing those who possess a product from taking advantage of those who need it. In an economic sense, it is a great waste of man power and materials to force nations to use substitutes when sufficient quantities of the real product can be produced at a lower cost than the substitute. Furthermore, the men and capital devoted to production of the natural product may have difficulty finding other employment except at great loss, and this is particularly true where capital takes the form of highly specialized productive facilities. A shift from a natural to a synthetic product, when the sole reason therefor is restriction on trade, may result in great loss to society.

We now are prepared to ask the question, Under what conditions should a good be produced at home, and when should it be imported? The principle involved is as follows: *Every community should call upon the resources of other communities if goods can be obtained at a cost that is lower than that which will be incurred in their production at home.* Again, *every community should specialize in the production of those commodities which it can produce at the lowest unit costs and exchange its surplus for the surpluses of goods produced in other regions.*

Principle of comparative advantage. The principle just stated has often been called the law of *comparative advantage.* As a formal law, it can be stated as follows: *A community, region, or country gains by specializing in the production of those goods in which it has the greatest*

comparative advantage. A corollary to this principle is as follows: *If a community, region, or nation does not possess a comparative advantage in any good, it should specialize in the production of that good in which it has the least comparative disadvantage.*

Let us discover how these principles operate and with what results. We shall first take a case in which each nation has a comparative advantage in one commodity, but each is producing both commodities.

Suppose that country A with a unit of labor and capital employed in the best combination can produce either 3 automobiles or 2 tons of cement. A second country, B, also employing its best combination can with a unit of labor and capital produce 2 automobiles or 3 tons of cement. If each country produces both commodities, production will be as follows:

Country A.	2 units *	3 automobiles + 2 tons of cement
Country B.	2 units *	2 automobiles + 3 tons of cement
Total		5 automobiles + 5 tons of cement

One unit of labor and capital used to produce automobiles and one to produce cement.

Since country A has a comparative advantage in the production of automobiles, let it specialize in the production of this good. Under these circumstances, it will employ 2 units of labor and capital in the production of automobiles and will produce no cement. The 2 units employed will in consequence produce 6 automobiles. Country B by specializing in the production of cement will produce 6 tons of cement. The result will be as follows:

Country A	2 units	6 automobiles + 0 tons of cement
Country B	2 units	0 automobiles + 6 tons of cement
Total.		6 automobiles + 6 tons of cement

Under these conditions, it is clear that each country has gained for, if they exchanged goods, country A could trade 3 automobiles for 3 tons of cement and country B would then have 3 automobiles and 3 tons of cement; whereas before it had only 2 automobiles. Country A gained by having an extra ton of cement, and country B had an extra automobile.

Suppose, however, that in the production of two other goods A has a comparative advantage in both commodities over B. What should be the policy of each country? Another example will aid us in determining the

principle to be used. Suppose that A by employing 3 units of capital can produce 60 bushels of wheat and 40 bushels of corn. Suppose that A devotes 1 unit to wheat and 2 to corn. B, however, is at a comparative disadvantage in the production of each of these goods. When B employs 1 unit in the production of wheat, it secures 40 bushels; whereas its 2 units devoted to corn produce only 20 bushels. If each country expends 3 units and produces both commodities, the results are as follows:

		1 *Unit*	2 *Units*
A	3 units	60 bushels of wheat	+ 40 bushels of corn
B	3 units	40 bushels of wheat	+ 20 bushels of corn
Total		100 bushels of wheat	+ 60 bushels of corn

Obviously A has a greater advantage over B in the production of corn than in wheat. If A specializes in the production of corn, since each unit expended will produce 20 bushels, when it employs 3 units, its total production will be 60 bushels. B, on the other hand, produces wheat. Since each unit expended produces 40 bushels of wheat, if all three units are employed, production will be increased to 120 bushels. The total production of each specializer will be as follows:

A	3 units	0 bushels of wheat	+ 60 bushels of corn
B	3 units	120 bushels of wheat	+ 0 bushels of corn
Total		120 bushels of wheat	+ 60 bushels of corn

By specializing, there has been a gain of 20 bushels in the production of wheat with the same expenditure of effort; whereas the quantity of corn produced has not decreased. How will this extra product be divided? Since country A can devote 1 unit of its labor and capital to the production of wheat and secure 60 bushels, its farmers will not produce corn for export unless they can get something more than 3 bushels of wheat for 1 bushel of corn. On the other hand, if country B shifted 1 unit of labor and capital to corn, it could produce 10 more bushels of corn at the cost of producing 40 bushels less of wheat. Therefore, the farmers of country B would not be willing to raise wheat for export unless they could get 1 bushel of corn for something less than 4 bushels of wheat. If the price of corn in terms of wheat is more than 3 bushels and less than 4 bushels, it will pay both countries to trade. It is difficult to say

which country will gain the more, because the price of each product will depend on supply and demand.

Let us suppose that the ratio of exchange is 3½ bushels of wheat to 1 bushel of corn. If B buys 20 bushels of corn, it will pay country A 70 bushels of wheat. The division will therefore be as follows:

A	70 bushels of wheat + 40 bushels of corn
B	50 bushels of wheat + 20 bushels of corn
Total	120 bushels of wheat + 60 bushels of corn

Owing to the new conditions, A will have made a clear gain of 10 bushels of wheat as will also B. Both countries gain in this exchange, even if A has a comparative advantage in the production of each of the commodities.

Modifications of the principle of comparative advantage. The application of the principle of comparative advantage has its limitations; otherwise, producers in each country might be able to specialize in the one commodity in which they had a relative advantage instead of producing hundreds or even thousands of commodities. The first limitation is based on the *principle of diminishing returns,* which we have already mentioned in connection with world trade. If producers in each country were to specialize in one commodity to the exclusion of all others, sooner or later they would encounter this principle; and the rising average and marginal costs as production increased would eventually offset the advantage formerly held.

Although the United States may possess a comparative advantage in production of automobiles, the law of diminishing returns is likely to operate and increase costs if the United States seeks to supply the needs of the whole world. Clearly the law of diminishing returns would come into operation if we tried to supply the world with wheat, and eventually any comparative advantage that we might have would disappear. In other words, the law of diminishing returns sets the limits beyond which specialization cannot go.

A second influence preventing absolute specialization is cost of transportation. Although goods may be produced more cheaply in one locality than elsewhere, the low-cost region may not be able to sell its goods in distant areas because of transportation costs. Hence, high-cost manufacturers who are nearer the market are able to compete if their initial

costs are no greater than the extra shipping charge that must be incurred by more efficient producers who are more remote from the market.

A third influence preventing absolute specialization has to do with difficulties in foreign exchange. We already have mentioned this consideration.

A fourth barrier to specialization arises from the fact that nations, either because of fear of war or to further domestic policy, impose barriers such as tariffs and quotas, or give subsidies or bounties to home producers who sell in foreign markets. Each of these devices prevents the flow of trade along natural channels and thus prevents or distorts specialization in world trade.

We should note that such barriers as the law of diminishing returns and cost of transportation may be classified as natural, since they will limit trade even though it is unhampered by governmental policy. The second two barriers, namely, interference with exchange rates and restrictions by tariffs and quotas, are artificial obstacles set up by government action. Later in the chapter, several explanations will be presented to account for governmental restrictions. At this point our purpose is only to call attention to the various forces that limit territorial and national specialization.

Mechanics of foreign trade. In our discussion of the mechanics of world trade we shall first deal with the factors that determine the rate of exchange when exchange is unhampered by governmental restrictions. The rate of exchange is the price paid in one country for a unit of the money of another country.

In this study we shall assume that two countries are on the gold standard and that no interference exists on the part of either government for the purpose of manipulating the rate. We must understand that the countries of the world are not on the gold standard at present; nevertheless the principles discovered in studying the gold standard are still useful to those who wish to understand foreign exchange.

If two countries are on the gold standard but employ different units as a basis for their monetary systems, each unit must be translated into a common measure. This can be accomplished by comparing the number of grains of gold in one unit with that in another. Prior to 1933 the British pound, for example, contained 113.0016 grains of fine gold, whereas the dollar of the United States contained only 23.22 grains of fine gold. Under these conditions the relationship between the pound

and the dollar was as follows: 113.0016/23.22 = 4.8664+. In other words, one British pound was 4.8664 times as valuable as one dollar, or was equal to $4.8664+. The relationship just explained represents the exchange value of the two currencies at mint par; that is, were no other costs involved or no other forces at work, one British pound prior to 1933 would have been worth $4.8664+. However, other forces are at work and these must be explained.

Bills of exchange. In order to understand the more elementary principles of foreign exchange and the mechanics of trade, we shall assume that payment takes place directly between an importer in the United States and an exporter in England. Under actual conditions, these men will deal through banks, but since this procedure is rather complicated, we shall simplify the problem at this point so that the principles can be established.

We shall assume that the importer in the United States places an order with an exporter in England. When the exporter receives the order, he proceeds to ship the goods. When the goods are sent, a bill of lading accompanies them. It may require that the transportation company collect for the goods before delivery, a method similar to our C.O.D. In this case, when the goods arrive in the United States, the importer pays the transportation company the proper amount. If, on the other hand, the importer and exporter have established a trade relationship, the situation is little different than when an individual establishes a charge account at a store, or when companies will allow 30 days for payment after goods arrive and are willing to accept checks. This practice is very common in the United States even though officials of the company are not personally acquainted with the purchaser.

Under the second assumption, when the goods are sent the exporter sends a bill to the importer, and upon the arrival of the goods the importer accepts them. Furthermore, the importer notifies the exporter that he has accepted the goods. He also writes "accepted" on the bill, signs it, and sends it back to the exporter. Thus the exporter has in his possession an instrument called an *acceptance.* Whatever the means of agreement, however, we call these instruments *bills of exchange,* and they originate with the exporter.

Gold points. Now what takes place when payment is made? The answer to this question depends in part on whether agreement has been made to pay in British or American funds. If payment is to be made in

American currency, the importer merely writes a check on his bank and the transaction is completed. The procedure is more complicated when payment must be made in British funds.

If payment is required in British currency, the importer seeks to purchase in the foreign exchange market the required funds payable in England. His own bank would be where he would go to acquire the British funds, since the large American and British banks maintain the market for foreign exchange. In other words, he would seek to pay American currency into his bank and to secure the ownership of an equivalent amount of British money in exchange.

Our familiar story of supply and demand now comes into the picture. If a number of American importers are seeking British funds, the price of British currency may be bid up sharply. However, under the gold standard there is a limit above which the price of British currency cannot rise. That limit is set by the cost of purchasing an equivalent amount of gold and shipping it to London. We use the term *gold exporting point* to designate this upper limit. It is the point above which the rate of exchange cannot rise, because beyond this point it becomes cheaper to export gold than to purchase British currency in order to make payment in London.

In the gold standard period, the expense of shipping gold from New York to London amounted to about 2½ cents per British pound. This covered the cost of crating, insurance, freight, and the like. Since the gold content of the two currencies made the British pound equivalent to $4.8664+ in U. S. funds, the gold exporting point roughly was 2½ cents higher, or approximately $4.89.

Thus, by way of summary, we can say that our United States importer would buy British currency in the foreign exchange market if he could purchase a British pound for less than $4.89. When the price of a British pound rose even a shade above this level, the importer would find it worth while to ship gold.

There also was a lower limit under the gold standard below which the price of British currency could not fall. This was $4.8664+ minus the 2½ cents' cost of transporting gold, or roughly $4.84. Just as it was cheaper for the American importer to acquire gold and ship it abroad if the price rose above $4.89, it would be cheaper for British importers to acquire gold and ship it to the United States if they could get no more than $4.84 for a British pound. If the rate of exchange fell below $4.84,

the British importers could buy gold in England, ship it to the United States, and get $4.8664+ for it. Even when they deducted the expense of shipping, the least they could net would be $4.84 for each British pound. Thus gold tended to flow into the United States if the rate of exchange fell even momentarily under $4.84. This $4.84 was the *gold importing point.* It was the lower limit below which the rate of exchange could not fall, because below this level gold began to flow into the United States.

Gold importing and exporting points between the United States and any foreign country on the gold standard were automatically set by the gold contents of their currencies and the cost of transporting gold between the two nations. Thus upper and lower limits were set to the fluctuations that might occur in the rates of exchange.

We have already implied that banks play an important role in handling the exchange of funds between nations. This role is independent of the existence of the gold standard. Today there is no international gold standard, but the role played by the banks in making foreign payments is as important as ever.

A large New York bank, for example, maintains deposits with a correspondent British bank in London or has its own branch there. When the New York bank is called upon in behalf of its customers to make payments in London in British funds, it can do so from the deposits that it maintains in London. In other words, the New York bank can accept funds in New York and pay them out in London because it has resources available in both cities. It replenishes its London deposits by taking in funds there from Britishers who have payments to make in United States currency. These payments are made from the New York resources of the American bank. As long as trade between the two nations is approximately in balance, the intake of funds in New York and London should approximately equal the outflow in both cities.

We can see that, under the gold standard, imports and exports of gold would provide a means for replenishing drains due to inequality of payments between nations. Such drains often were due, of course, to lack of balance in trade between two nations. When no gold standard exists, the laws of supply and demand for a foreign currency may serve as a brake to a rise or fall in rates of exchange. For example, if importers must pay a very high price for foreign currency in order to purchase goods, they are reluctant to import. As a result, the demand for the

foreign currency declines and the upward pressure on its price is reduced. Earlier in this chapter we noted, however, that the absence of the gold standard tends to increase greatly the fluctuations in rates of exchange, except when these are pegged by government edict, or through the working of a stabilization fund that offsets fluctuations by the purchase or sale of foreign currencies.

Buyers and sellers of bills of exchange. Many thousands of transactions take place every year between businessmen residing in different countries. Exporters have bills of exchange for sale and importers wish to buy them. The fact that there are many exporters and importers in every country makes a market in bills of exchange possible. This market is like any other in that price is determined by supply and demand. Under the gold standard, the price for bills of exchange could fluctuate only between the gold importing and exporting points. As we saw in the preceding section, the law of supply and demand would impose limitations in normal times to extreme fluctuations in bills. We saw also that governments sometimes act to avert extreme fluctuations.

Indirect exchange. Few nations have even close balance between imports and exports with each individual country. Such balance is not necessary. The important consideration in regard to the trade of a nation is for the sum of exports to all other countries to be equal to the sum of imports from all other countries. Often an excess of imports from a particular country can be offset by an excess of exports to some other nation.

Suppose that the United States sends to England more goods than are purchased from England. However, the United States buys large amounts of coffee from Brazil and does not sell as much to Brazil as it buys. England, on the other hand, sells more goods to Brazil than she buys from this country. Bankers in these three countries could balance the exchanges by using the credit that England has built up in Brazil to offset the debt that England has incurred in the United States. Such indirect exchange is quite common and makes for a more efficient world trade.

Favorable and unfavorable balance of trade. Since much is made of the fact that nations should have a favorable balance of trade, we shall define this term. Many services or transactions take place between countries which are not recorded as commodity trade. Thus, although our merchandise exports may be greater than our merchandise imports, this condition is but one of many, though generally by far the most im-

portant, that must be taken into account in determining whether a nation has a favorable or an unfavorable balance of trade. Until the latter part of the eighteenth century the theory prevailed, as we have pointed out earlier, that nations should encourage exportation of goods and discourage importation. This theory has been called "mercantilism." From this theory arose the idea that, if more merchandise in foreign trade was sold than purchased, the country had a *favorable balance* of trade. On the other hand, if merchandise imports exceeded exports, the country had an *unfavorable balance.* This philosophy, as we have said, still has many adherents.

Visible and invisible items in trade. Most governments record rather carefully the quantity of merchandise and gold and silver that is exported and imported. Since these items are material, they are called "visible" items and are used to determine whether or not foreign trade is in a favorable or an unfavorable position. However, much business is transacted between nations aside from such visible items. These transactions include such things as freight charges, tourist expenditures, immigrant remittances, interest payments on investments, and royalties. These are called "invisible" items, and a favorable invisible balance may offset an unfavorable visible balance of trade. For many years, Italian immigrants in this country sent money to their relatives in Italy. In sending this money, they really purchased bills of exchange from American producers who were exporting goods to Italy. Importers of such goods in Italy would pay for goods purchased in the United States by purchasing these immigrant remittances from individuals in Italy. Thus, no visible item had to be sent to Italy from the United States to take care of these remittances, nor did gold have to be sent from Italy to the United States to pay for goods imported, insofar as the amounts balanced. England had an unfavorable visible balance of trade for years before the Second World War, but she also had vast investments in foreign countries. Interest, royalties, and similar invisible payments that foreign countries owed to Englishmen offset the unfavorable visible balance; hence, little or no gold was needed to balance payments for goods imported into England.

In considering the balance of trade, therefore, we must note invisible as well as visible items; and what may seem to be an unfavorable balance at first sight may actually be a favorable balance if all items in trade are considered.

Our discussion of the nature of exchange has been concerned with the simple mechanics involved in making payments when two or more countries are on the gold standard. We have already pointed out that the nations of the world are not now on the gold standard and that exchange rates are often manipulated by governments for various purposes. In our study of the principles underlying foreign exchange these problems cannot be discussed. We need only to note that such manipulation will vitally affect trade. *A nation must pay for the goods and services that it buys with the goods and services that it sells* is a truth that still holds good. The same principles apply to nations as to local communities; the only difference is that international trade becomes complicated by the necessity of changing domestic money into foreign money and by the existence of special trade regulations.

Why trade between countries must balance in the long run. The analysis to this point has been concerned with the methods by which payments are made in foreign trade. We know that in the long run imports must equal exports, for any country can pay for the goods that it buys only with the goods that it sells. We must now explain why in the long run trade between countries must balance.

Let us assume that two countries engage in trade. We may ask, Why should one country purchase goods in another country rather than purchase goods at home? Our study of the principle of comparative advantage gave one answer to this question. There is another answer, however, which must be taken into account. Suppose consumers in country A discover that prices in country B are lower than those which prevail at home. Since there is free trade, consumers will purchase in country B. In other words, exports from country B will increase. Consumers in country B, however, will not purchase goods in country A because prices are higher than those which prevail at home. Under these conditions, business will increase in country B and decrease in country A. Under free trade and competition the demand and supply will operate, with the result that prices will rise in B while they will fall in A. The result will be that soon prices in both countries will come into equilibrium and trade will no longer flow in one direction only. Under a system in which exchange is unhampered, therefore, trade may be carried on as between nations in a similar manner and with the same benefits as it is at home. Only when governments interfere with the automatic operation of this system do consumers fail to secure the full benefits of world trade.

Failure to understand the basic principle of the nature of world trade has caused many countries to impose restrictions in order to create a *favorable* balance of trade. Such attempts prevent nations from producing goods and services in which they have a comparative advantage and from exchanging these goods and services for those of other nations in which the latter have a comparative advantage. Consequently, nations and consumers are deprived of goods and services that they might have enjoyed were no barriers imposed.

Foreign trade with nations off the gold standard. Our study of trade under the gold standard has been useful in learning the principles governing the exchange of goods and services between nations. However, the nations of the world no longer are on the gold standard. What about trade in a world that has forsaken the gold standard?

As long as a nation's payments abroad equal its receipts from abroad, there should be no difficulty in carrying on a heavy volume of trade without the gold standard. In other words, the essential condition, with or without the gold standard, is that the sum of invisible and visible exports must equal the total of invisible and visible imports. An excess of export items will result in an accumulation of foreign balances abroad, which is the equivalent to lending to foreigners. If more is imported than is exported, the balances of foreigners in the home country will be built up. This is equivalent to borrowing from foreigners.

Such disequilibrium cannot continue forever. Under the gold standard inflows and outflows of gold automatically would bring about an adjustment through price level changes or by forcing changes in rates of exchange, or perhaps both. Without the gold standard, artificial arrangements must be made to bring trade into balance. Adjustments of exchange rates under auspices of the International Monetary Fund may bring temporary relief. One of the functions of the fund also is to confer with nations on maladjustments and help correct them. The resources of the World Bank also may be made available to nations with adverse balances of trade, so that they may build up their industries and develop their resources until they can win a more secure place in trade.

The essential point is that a flourishing world trade can be carried on without the gold standard. If the International Monetary Fund and the World Bank fulfill the hopes of its members, the instability of exchange rates which plagued trade during the 1930's need not be repeated. Goods and services can move across international boundaries with confidence

that payment in stable currencies will be forthcoming. The usual commercial channels for transfer of funds will continue to function with the fund ready to step in during periods of unusual strain.

Trade restrictions. We have pointed out above that nations, besides manipulating foreign exchange and the movement of gold, often resort to other types of artificial trade restrictions, such as tariffs, quotas, and subsidies. We may now ask the question, If trade that is free from restrictions is beneficial to all concerned, why are trade barriers imposed? The answer is extremely complicated, for it encompasses all those forces which, as pointed out above, are the essence of nationalism. We shall, however, discuss a few of these restrictions.

Valid arguments in favor of imposing trade barriers. It is perhaps natural that every nation desires to become self-sufficient. Fear of war, desire to foster and develop infant industries, the desire to secure diversity in order to have a better balanced economy are all worthy motives and have a strong appeal to the citizens within a country. To the extent that these motives are valid, barriers may be justified. Let us, therefore, review each of the arguments briefly.

1. Fear of War. Every nation and every generation deplores the destruction and waste that accompany war and the preparation for war. As yet, few nations have been able to escape the fear of war. As long as the possibility of war exists, nations must develop and maintain certain key industries that can be called upon to produce weapons in time of war. In the face of possible wars, it is folly for nations to fail to develop these key industries. Thus, tariffs or other restrictions may be imposed upon products of other nations that may seek to compete with these industries for the trade of the country. The problem here is to determine which industries are key industries. Often every industry wishes to qualify as a key industry and secure by this means partial immunity from world competition.

2. Infant Industries. A second valid argument for the imposition of barriers concerns infant industries. When countries are new or backward industrially, they may possess the potential resources to produce a commodity. However, time is needed to start an industry, and costs in the beginning may be relatively high. Foreign nations that are more highly industrialized can produce the goods at lower prices. Hence, under free competition, the new industry will not be able to secure a market. In order to encourage new industries, therefore, barriers are placed on

foreign trade so far as this product is concerned. These barriers can be justified if the infant industry that is being aided can show that in a reasonable future it will be able to produce the product at a unit cost as low as or even lower than that in foreign countries. If the barrier is temporary and designed only to allow the industry to secure a foothold, then there may be a gain in the long run. The danger is that some of these industries may never become very efficient, and that others will seek to profit unduly or even secure a monopoly while demanding continued protection.

3. Better Balanced Economy. A third valid argument is that industry in each nation should be sufficiently diversified in order to have a balanced economy. In order to secure such diversification, some industries may be protected. The argument here is that specialization may lead to maladjustment. Thus a nation, like a city, should not be devoted entirely to the production of one product. If labor is used for but one product, undoubtedly many difficulties would result. During periods of prosperity, for example, such a nation might seem to profit unduly, but the depression period might result in absolute stagnation. Wisdom favors the policy of balancing any economy so that it will be partly industrial and partly agricultural. Industrial centers are better able to withstand booms and depressions if they possess a variety of industries. The movement in this country leading to decentralization of industry in order to secure diversification and to prevent too much overspecialization among workers is a movement that recognizes these dangers. The problem so far as foreign trade is concerned is again to determine to what extent it is necessary to penalize foreign goods in order to secure reasonable diversification. In a country like the United States there is sure to be extensive diversification without employing artificial measures to secure it.

Invalid arguments in favor of imposing trade barriers. We have pointed out that tariffs, quotas, or subsidies may be justified under certain conditions. Although these practices represent artificial interferences with world trade, in the long run the benefits received may be worth the extra cost entailed. However, certain arguments advanced in favor of imposing barriers cannot be justified.

1. Home Market. Thus, it is insisted that barriers should be imposed in order to create a home market. We have pointed out the conditions under which such an argument is valid. Beyond these limits, the demand that protection be given to home industries violates the fundamental

basis of all trade, and in consequence the consumers of the world will have fewer goods. Although an increase in business at home may seem to be a net increase, it represents merely a shift, and a poor shift at that, for when the shift in industry is made, those who would be more productive, if world trade were carried on, are now less productive. The earlier discussion on the relative merits of specialization shows why a tariff barrier whose purpose it is to enlarge the home market causes only a shift from a more efficient to a less efficient utilization of the resources of the country. Moreover, such a policy will result in a loss of total trade. The "Buy it at home" and "Patronize home trade" slogans cannot be justified except within the limits discussed above.

2. STANDARD OF LIVING. A second argument to be considered is often called the *standard of living* argument. Proponents maintain that barriers must be erected to protect the wage rates of American labor against low-paid foreign workers. This argument is also advanced when a clamor arises for lower tariffs. Here the claim is made that high wages are paid in the United States because the American worker is protected by the tariff.

The most telling argument against this point of view is to point out that the ability to compete in foreign trade does not depend upon low wages, but rather upon low unit costs. Wages are relatively high in the United States; yet unit costs are often extremely low. Wage rates are based in the main, as will be shown in a subsequent chapter, on productivity. When productivity is great, high wages can be paid. The problem is to secure the best combination of factors; and when this is secured, low unit costs result, even though wages or payments for the use of land or capital are relatively high. Thus, emphasis must be placed on unit costs, not on wages. A second answer to the standard of living argument may be given by resorting to statistics. No reliable figures have been found to prove that wages are higher in industries that are protected than in those not protected. Furthermore, the record does not show that, when industries received aid in the form of tariffs or quotas, they increased wages above those paid in other industries. Wage rates and the standard of living depend upon factors other than trade barriers. Unit costs are the determining factors in trade. In short, wages are high in the United States because the productivity of American workers is high. A third answer to the standard of living argument is that any increase in

wages made possible by a tariff is ordinarily paid by American consumers in higher prices and this reduces the purchasing power of all American incomes.

Summary

This chapter has pointed out that world trade, like trade at home, results from differences in climate and in the distribution of resources among different countries. Because of these differences, producers in some communities or countries have a comparative advantage over others. By producing those goods which can be manufactured at low costs and purchasing those which others can manufacture cheaper, all persons benefit through exchange.

The principle of specialization, however, has definite limits. Transportation costs and the law of diminishing returns ultimately increase costs, and consequently producers with higher unit costs can continue to bring goods to the market. This principle is further limited because of the fear of war, the desire to aid infant industries, and the desire to maintain a better balanced economy. In order to enjoy these latter advantages, artificial barriers are imposed. These are of two kinds: manipulation of exchange rates and the use of tariffs, quotas, and similar barriers. The danger in the use of these barriers lies in the fact that they disturb world trade and nullify the economic advantages that might otherwise accrue to all engaged in world trade.

Study Questions

1. Why does a flourishing foreign trade seem desirable for a nation? If in your opinion a flourishing foreign trade is not desirable, defend your position.
2. What are some of the natural barriers to foreign trade? What would you say was the principal artificial barrier to foreign trade? Can you think of other artificial barriers?
3. What is the principle of comparative advantage? Illustrate it by an example. What modifications of the principle can you enumerate?
4. What is the basis under the gold standard for establishment of rates of exchange between two currencies? Can you think of how rates of exchange could be established between two countries not on the gold standard? What about rates of exchange between a gold standard country and one not on the gold standard?
5. What is a bill of exchange and what is its place in the financing of foreign trade?

6. What is the popular meaning of the terms favorable and unfavorable balance of trade? Can you think of any circumstances under which it might be advantageous for a country to have an unfavorable balance of trade? How would you classify the position of the United States at the present time in regard to the desirability of an unfavorable balance of trade?

7. Distinguish between a visible and an invisible item in foreign trade. Give some examples of each.

8. What are some of the arguments frequently made in favor of trade barriers? How valid do you regard each argument? Defend your position.

9. Under the old gold standard, what role did gold play in foreign trade? What do we mean by gold points? If a gold flow started from one nation to another, what might halt it and perhaps eventually reverse it?

10. In view of what you have learned in this chapter, what trade policy, especially in relation to tariffs, would you advocate for the United States at the present time?

Exercises

1. Suppose that Worker A can produce 30 units of X good in a day and with the same effort produce only 15 units of Y good. Worker B, on the other hand, in the same period of time can produce 15 units of X good and 10 units of Y good.

Would A and B both specialize and, if so, would specialization be advantageous to both? Why or why not? In answering this question consider the following:

a. Does A have an absolute or only a comparative advantage over B?

b. If A and B each worked 1 day producing X good and 1 day producing Y good, what would be the total output of each good?

c. Suppose that A worked 2 days producing X goods and B worked 2 days producing Y goods, how much would be produced by both of them?

d. Suppose that A can secure $6 a day for his labor while B can secure only $4 a day. In terms of cost, how much is the labor cost in producing a unit of X goods and Y goods?

e. Would A be willing to pay B somewhat more than A's cost for Y goods, if he were allowed to specialize in X goods? Why or why not?

f. Suppose that the price of X goods is 25 cents per unit and Y goods is 42 cents per unit. Would A and B be better off by specializing in one good?

Why? What would be the income of each per day under the above conditions? Are both better off by resorting to specialization?

g. Suppose that the price remained at exactly what it cost A to produce each good. Would A still be better off to specialize in X good and allow B to produce Y good? Why or why not?

h. Why is it unlikely that the price of these commodities would remain at A's costs?

2. If we buy goods abroad, we get goods but foreigners get the money; if we buy goods at home, we get both the goods and the money. Is this a valid argument?

3. If Canada and the United States were one country, would this change our trade relations? Which country would be harmed if trade were free as between these two countries? Why?

4. Suppose that Illinois decided that too many of her citizens were spending their money in Florida. She therefore decided to restrict travel to Florida and to raise oranges and grapefruit in Illinois. In what way would Illinois be harmed by this procedure? How would Florida be harmed? Would not the same principles apply as between Illinois and Mexico, England, Australia, or any other country? In the light of your answer defend trade barriers in foreign trade.

5. The United States is now a great creditor nation. In respect to foreign trade should a creditor nation have a favorable or an unfavorable balance of trade? Explain.

6. Certain individuals maintain that in normal times 90 per cent of our trade is domestic; therefore, we do not have to worry about foreign trade. Should we be concerned with foreign trade if that statement is correct? Why or why not?

7. It is maintained that, unless we engage in foreign trade, we cannot use our agricultural or industrial plant to capacity. But if we sell, we must buy. List some of the outstanding imports that could be used in this country in exchange for our exports.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chaps. 26 and 27.

BYE, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chaps. 17 and 18.

DEIBLER, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chaps. 15 and 16.

FAIRCHILD, F. R., E. S. FURNISS, and N. S. BUCK: "Economics," The Macmillan Company, New York, 1940, Chaps. 31-34.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 23 and 24.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chaps. 33-37.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chaps. 23 and 24.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chaps. 24 and 25.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chaps. 24-26.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. I, Chaps. 31-37.

20. Distribution

Meaning of distribution. The division of the fruits of production among the factors that contribute to the productive process, or among the people who control these factors, is what economists call *distribution.* Distribution in this sense should not be confused with *marketing.* Marketing is a study of the physical and economic processes by which particular goods are transferred from the producer to the ultimate consumer. It includes the study of such activities as transportation, storage, and merchandising. Distribution, on the other hand, is concerned only with the amount of income that goes to the different factors of production or to those who control them.

Whenever goods are sold in the market, the entire price paid for them ultimately represents income to someone. This does not mean, of course, that, when a manufacturer receives $1,000 for an automobile that he has produced, the entire $1,000 is net income to him. Most of the price that he receives merely reimburses him for costs that he has incurred. However, all the payments and allowances that the manufacturer regards as costs represent income to someone else, either directly or indirectly. Ultimately most of the price received for a good goes to workers in the form of *wages,* to landowners in the form of *rent,* and to the owners of capital in the form of *interest.* If something is left over, this goes to the businessman himself and is called *profit.*

Some readers may question the statement that the entire price a manufacturer receives for his product goes to pay rent, wages, interest, and profits. They may, for example, recall that raw materials are an important cost of production for most manufacturers. However, prices paid for raw materials serve merely to meet wage, interest, and rent costs incurred earlier in the productive process, or to provide profits for other businessmen. This is true not only in manufacturing but in all productive enterprises. Even the cost of depreciation on a building is a charge made to offset other payments that were made to land, labor, and capital when

the building was constructed. It is, therefore, substantially correct to say that the entire value of the economic goods produced by society is received as income by those who control land, labor, capital, or business enterprise.

To reduce the problem of distribution to its simplest terms we might think of the total product of society as a pie that is to be divided into four pieces. Each of the four factors of production, land, labor, capital, and entrepreneurship, contributes to the preparation of the pie and each will receive a slice. Our chief problem is to find how the size of each share is determined.

Functional distribution and personal distribution. *Functional distribution* is the study of the division of income among the four factors of production, regardless of who controls them. *Personal distribution* is the study of how income is divided among the persons who make up the community.

In a free-enterprise economy, functional distribution is merely a branch of price theory. The share of the product that a factor receives is the price paid for the use of the factor; and the chief problem is to find how this price is determined. Wages is the price of labor; rent is the price for the services of land; and interest is the price for the services of capital. However, to what extent profits should be regarded as the price for the services of the entrepreneur is open to some question and is a problem that will be dealt with in the chapter on Profits.

The problem of personal distribution is somewhat different. In order to understand personal distribution, or the division of income among individuals, we need, of course, to understand how income is divided among the factors of production. In addition, however, we need to understand why there is such an uneven division of land and capital among the people who make up our society. In a free society everyone controls his own labor and can receive in wages whatever his labor is worth in the market. Many people, however, possess no land or capital; at the other extreme, a few control so much land and capital that they can receive very large incomes in the form of rent and interest, and sometimes in the form of profits.

In order to understand how the present unequal distribution of land and capital came about we should have to study the history of social institutions. Of special importance in this connection would be the laws and customs regulating the ownership, inheritance, and taxation of pri-

vate property. It is clear that the personal distribution of wealth and income is a matter of great social importance. It is not, however, a problem that we shall attempt to deal with to any extent in this book. Our principal concern here is with the functional distribution of income, and we shall treat functional distribution as a problem in price determination. However, before we turn our attention to this problem it might be well to take note of certain important questions that arise in connection with the personal distribution of income.

The fact that very large incomes are received by a few individuals has caused frequent demands for a "more equitable" distribution. Some believe that there is something wrong per se in permitting the enormous incomes that have in the past accrued to certain people. Yet there are some social advantages in this unequal distribution of income. For example, persons with large incomes are best able to accumulate the great stores of capital that a progressive society must have. Also, educational, charitable, philanthropic, and cultural institutions depend upon the contributions of persons with large incomes. Moreover, it is easy to show that the leveling of such incomes would not do much to raise the general well-being, because the number of large incomes is so small in relation to the total number of income receivers.

In some nations aggressive efforts have been made to limit the size of the income that any individual may have. That is the effect, for example, of the highly progressive income tax rates in the United States, since these rates increase rapidly as income rises. However, there always is the danger under such a tax system that rates will be pushed so high that capable individuals with capital will lack adequate incentive for establishing new business enterprises, or expanding old ones. Of course such expansion is very important if sufficient employment opportunities are to be provided and if standards of living are to continue to rise. Insofar as large incomes are a necessary incentive to productive effort, they perform an essential social function.

Prices of the factors of production. Since payments received by factors of production are for the most part simply prices, it is possible to apply to them the general principles of price determination. However, the prices of the factors of production present some unique problems of considerable importance, and to these it will be necessary to give special attention. Such special problems are the principal justification for this

chapter on Distribution, and for the chapters which follow it on Rent, Wages, Interest, and Profits.

Productivity as the source of the demand for a factor. The demand for consumers' goods arises from their utility. People are willing to buy things like eggs and shoes because these goods directly satisfy personal wants. However, land, labor, and capital used in production do not directly satisfy wants. Hence the demand for them is commonly said to be derived from the demand for their products. More directly, the demand for a factor of production is derived from the *value* of its product. For example, the real reason that the owner of a shoe factory is willing to hire a certain worker at $10 a day is that he believes that the worker will add at least $10 to the net value of the product of the factory.

However, this raises the question of how the manufacturer can separate the product of one worker from that of another. Many workers contribute to the making of a shoe. Is it possible to determine what part of the shoe or what part of its value was produced by any one worker? Furthermore, it is not only workers that contribute to the making of the shoe, but also capital, natural resources, and management. Is there any way of determining what part of the shoe, or what part of its value, should be attributed to each of the major factors of production, land, labor, capital, and entrepreneurship?

It is clear that in an absolute physical sense there is no way of separating the product of a shoe factory into the output of land, the output of capital, the output of labor, and the output of management. Neither, in an absolute sense, is there any way of determining which factor of production is most important, for ordinarily, unless all the factors cooperate, there will be no product at all. To ask which factor is most important is a good deal like asking a man who is hanging from a cliff by a chain which link of the chain is most important in keeping him from hurtling to his death in the chasm below. Yet if we are to explain the demand for a particular factor of production in terms of the value of its product, we must in some sense be able to separate the product of one factor from that of another.

One way to solve this problem would be simply to assume that a factor of production contributed to the value of the product an amount equal to the cost of the factor. For example, a worker who is paid $10 a day in a shoe factory would be arbitrarily credited with a contribution of $10 to the value of the shoes. For our purposes, however, this solution

of the problem is not helpful, because what we wish to determine is *why* the worker receives a wage of $10.

In a competitive market, wages, like other prices, are determined by supply and demand. If we are to explain the demand for labor as being derived from labor's contribution to the value of the product, we must determine labor's contribution in a way that is not directly dependent on the wage rate if we are to avoid a naïve type of circular reasoning. To help solve this problem of determining how much labor, or for that matter any other factor, contributes to the total product, economists have resorted to the concept of *marginal productivity.* This concept will be explained in the paragraphs that follow.

Marginal productivity. To simplify the problem of explaining marginal productivity it will be assumed that all units of a given factor are similar in every respect, so that any unit may be substituted for any other without loss of efficiency. It will be recalled that in our study of the law of diminishing returns we applied units of a variable factor to a fixed factor and measured the resulting product. Throughout this process we focused our attention on the addition to the total product resulting from the use of the last unit of the variable factor applied. This last unit of the factor, as we noted, is called the "marginal unit"; and the increase in total product resulting from its use is known as the "marginal product." For convenience we shall use the term *marginal productivity* to refer to the price of the marginal product of a factor. We can then say that, the other factors of production being held constant, *the marginal productivity of any factor of production is the price of that part of the total product which depends on the employment of any one unit of the factor.*

Since marginal productivity is the price of the amount of product that depends on any one unit of a factor of production, it is also the price of the amount of product that depends on each and every unit of the factor. The student will recall that we have assumed that all units are similar. Only at the margin is it possible to discover and separate out the specific product of each factor of production. In a given situation, the marginal productivity of a factor represents the maximum amount that any producer would be willing to pay for any one unit. For example, if a day's labor by a worker would add $10 to the net value of a manufacturer's product, $10 would be the most that he could afford to pay for this labor.

Marginal productivity schedule. We know from the law of diminishing returns that beyond a certain point the marginal product of a factor will

decrease as more units of it are employed. As a result, the marginal productivity of the factor will also decrease, provided there is no change in the price of the product. This will be true for a given enterprise, and it will also tend to be true for the economy as a whole; always assuming, of course, that the quantities of the other factors remain fixed. If a table could be drawn up showing the relation between the quantity of a factor employed and the marginal productivity of the factor, such a table would be a *marginal productivity schedule.*

Table 26.—Variations in the Marginal Productivity of Labor with Variations in the Quantities Employed

Number of Workers Employed Eight Hours Daily	*Marginal Productivity of Labor*
35,000,000	$8.00
40,000,000	7.50
45,000,000	7.00
50,000,000	6.50
55,000,000	6.00
60,000,000	5.50
65,000,000	5.00
70,000,000	4.50
75,000,000	4.00

Since economists are generally agreed that under actual economic conditions each factor is employed well beyond the point of decreasing returns, we need concern ourselves only with the declining phase of a marginal productivity schedule.

Let us use labor to illustrate a marginal productivity schedule. We shall assume the employment in an economy of various amounts of labor without any change in the amounts of land and capital available. Further, we shall assume that there is no change in the price level of the products of industry,[1] and that therefore a given change in the physical output of goods will represent a corresponding change in value output. Under these conditions, if relatively small amounts of labor are employed, mar-

[1] To make this assumption tenable it is also necessary to assume that the money supply is varied in such a way as to stabilize the price level of output. Otherwise changes in the amount of labor employed would change the price level of goods both by changing the national income and by changing the quantities of goods produced and offered on the market.

ginal productivity will be comparatively high; but as larger amounts of labor are used, marginal productivity will decline. To make the illustration as simple as possible we shall also assume that all workers are exactly alike, and that the labor unit is an 8-hour day.[2] Table 26 represents this situation. Figure 1 is a graphical representation of Table 26.

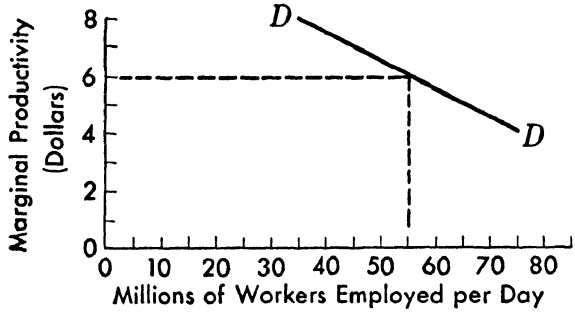


Fig. 1.—Demand for labor as indicated by marginal productivity.

Market demand for a factor of production. A schedule showing marginal productivity when various total amounts of a factor of production are employed is also a market demand schedule for that factor. This is true because, as we shall see, it can be demonstrated that at any given wage rate employers will hire more men up to, but only up to, the point where the marginal productivity of labor is equal to the wages paid. It follows that Table 26 and Fig. 1 represent not only the marginal productivity schedule for labor but also the demand schedule. Figure 1 shows, for example, that at a wage rate of $6 per day 55,000,000 workers will be employed. It also shows that when 55,000,000 workers are employed the marginal productivity of labor will be $6.

Let us see why employers will tend to hire just enough workers to make the marginal productivity of labor equal to the wage rate. Suppose, for the moment, that marginal productivity and the wage rate are not equal. Assume, for example, that the wage rate is $6 but that only

[2] Workers of course are not alike. They differ in the kinds of work they can do, the degree of skill they possess, and the amount of product they turn out per hour. These differences, however, do not invalidate the principle being illustrated. Even though labor is not a uniform commodity, if the total amount were increased, including all kinds and grades, the average marginal productivity of a day's labor would fall.

50,000,000 workers are employed. With 50,000,000 workers employed Fig. 1 shows that the marginal productivity of labor will be $6.50. This is clearly not an equilibrium situation, for it will pay employers to take on more workers. Although an additional worker will cost the employer only $6 a day, he will increase the value of the product by $6.50. It is clear that, as long as the marginal productivity of labor is above the wage rate, employers can increase their net profits by taking on more men. As soon, however, as marginal productivity equals wages, there is no longer an incentive to hire more workers. Furthermore, if marginal productivity drops below the wage rate, it is actually to an employer's advantage to reduce his working force.

It is clear, then, that in a freely competitive market the demand schedule for labor tends to be identical with the marginal productivity schedule. The same principle applies to all the factors. The demand for a factor is derived from its productivity, and the demand schedule for any factor is identical with its marginal productivity schedule.

Market supply of a factor of production. The supply of a factor of production is a schedule of the quantities of it which would be offered for use at various prices in a given market and in a given period of time.[3]

Strictly speaking, it is not the factor that is offered for sale but its services or use. Since the services of some units of a factor may be withheld from the market, especially at low prices, the supply of a factor is not the same thing as the quantity in existence. However, the quantity in existence is the most important single element in limiting supply at any given time.

One peculiarity of the factors of production is that their prices, or the prices of their services, cannot be related to their cost of production in

[3] Supply and demand schedules, as has been noted earlier, are static concepts. In constructing them one assumes that for the period in question all elements in the economy remain constant except the price of the commodity and the amounts offered or demanded. Strictly speaking, this assumption is untenable, for theoretically a change in the price or in the quantity exchanged of any given commodity is likely to affect the prices and quantities exchanged of all other commodities. In the case, however, of commodities like corn, aluminum, or shoes, the assumption that other things remain constant may not be very much in error. The error is considerably greater when the prices and quantities under discussion are those of the factors of production. It would hardly be possible for wage rates, interest rates, or rents to change very much without appreciable effects on almost everything else in the economy.

any significant sense. Land is a costless gift of nature. Labor is human beings or their services, and nobody has yet found a satisfactory way to determine the money cost of producing either a man or a day's labor. Capital goods, of course, do have a cost of production, and a price that tends to equal this cost. However, in the study of distribution our concern is not with the price of a capital good per se, but rather with the price for the *use* of capital, that is, with the interest rate; and interest in the economic sense is a net payment over and above any production or depreciation costs on capital.

The available quantities of the different factors are limited in somewhat different ways. Since land is natural resources, it is rather definitely limited in total amount, and the limitation is a permanent one. The situation with respect to labor is not quite the same. At any given time the potential supply of labor is also limited, for it depends on the size of the population, but over a period of time the population changes. However, the forces that cause population to change are not very well understood. It might be supposed that high wages would bring an increase in population by inducing people to have more children; but in fact high wages seem to have the opposite effect. The source of capital is saving. In any short period of time the total amount of capital in the economy is approximately constant. However, in a dynamic economy like ours the quantity of capital tends gradually to increase as a result of saving; and it is possible, though not certain, that the interest rate is a significant factor in inducing people to save.

Determination of the price or share of a factor. In a competitive market the price of the services of a factor is determined by demand and supply at the point where the amount demanded is equal to the amount offered. At the equilibrium price there can be no unemployment of the factor. Everyone who is willing to sell his labor or lend his capital or rent his land can do so at the market price. Likewise everyone who wishes the services of labor, land, or capital can obtain them at the market price. Unemployment of the factors of production, as for example idle labor, is a result of limitations on competition. Since in the actual markets for the factors of production these limitations are of great importance, considerable attention will be given to them in the chapters that follow.

Using labor again as an example, Fig. 2 illustrates the determination of the price of a factor of production. SS represents the supply schedule for

labor on the assumption that workers will offer somewhat more of their services at high wage rates than at low ones. *DD* represents the demand schedule for labor and also the marginal productivity schedule. At equilibrium both the wage rate and the marginal productivity of labor are represented by *PR;* and the amount of labor employed is represented by the distance from the origin to R.

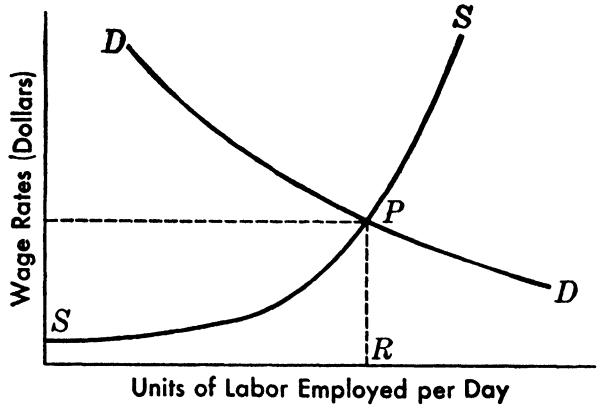


Fig. 2.—Demand and supply curves for labor in a given market in a given short period of time.

It should be emphasized that, though the price of a factor tends to equal its marginal productivity, marginal productivity does not determine price. Rather, both price and marginal productivity are determined simultaneously. They are the resultant of all the forces that are operating in the market to influence either supply or demand.

It is easy to show that marginal productivity is not a fixed or independent quantity, but that it depends on a number of circumstances. Some of these are worth noting: (1) The marginal productivity of a factor depends partly on the quantity of the factor employed, because the marginal product is subject to the law of diminishing returns. (2) The marginal productivity of a factor depends partly on the available quantities of the other factors. For example, if all laborers receive more and better machines to work with, the amount of product dependent on any one unit of labor will rise. (3) Marginal productivity also depends on the price of the product. For example, if the price of wheat rises because of an increase in demand, the value of the marginal product of labor in the wheat fields will also rise. This of course will cause an in-

crease in the demand for labor and some rise in wages. The increased employment of labor will, in turn, result in some decline of marginal productivity from the new high level. When all these adjustments, and others, have been made, equilibrium will be reestablished.

Prices of unlike units of a factor of production. So far we have assumed that all units of a factor of production were exactly alike. It was necessary to make this assumption in order to explain in a fairly simple manner the nature of the forces that determine the price of a factor in a competitive market. The assumption, however, was contrary to fact, for in reality different units of a factor of production are seldom alike. Often, indeed, they are so unlike that they may even be regarded as different commodities. However, the existence of differences between the units of the factors of production does not invalidate our theory of price determination.

Even though the units of a factor of production may be quite unlike one another in important respects, in a competitive market it is still true that the average price level for all units of the factor, and also the average level of productivity at the margin, will be determined by the forces of supply and demand. In this case, however, the prices of different units will not be the same; rather, the price of each individual unit will be determined by its marginal productivity. It is clear, for example, that if one worker would increase the value of a farmer's product by $5 a day, while another and more efficient worker would increase it by $10 a day, the farmer would not be willing to pay the first worker more than $5, while he would be willing to pay the second one anything up to $10. In a highly competitive labor market, the competition of other farmers for workers would tend to force him to pay to either worker the full value of the marginal product.

Largely because different units of a factor of production are different, the *contractual* price or share of a factor often differs from the economic price. For example, if all units of land were exactly alike in every respect, the law of one price would apply, and at any given time a single market rental price would prevail for all units. This price, of course, would be equal to the marginal productivity of land, and would therefore represent the true economic rent. Since, however, land varies greatly in productivity, the marginal productivity of each piece of land must be estimated separately. As a result, both owners and renters may

make numerous errors, so that the rent agreed upon by contract may be either more or less than the true economic rent.

The fact that the units of a factor of production are often unlike has still another effect, and that is to introduce elements of monopoly into the market for the factor. Take labor as an illustration. No organized group can possibly have a monopoly of labor *in general.* However, different kinds of labor are really different commodities; and it is quite possible for a strong union, say the plumbers' union in a certain large city, to exercise monopoly control over the supply of a particular kind of labor. This problem will be discussed more fully in the chapter on Labor.

Simultaneous determination of wages, interest, rent, and profits. It should be emphasized that the price of a factor of production cannot possibly be determined *independently* of the prices of the other factors. In earlier times economists often tried to do this. They hoped to find some one simple explanation for the price of each factor. This is impossible because the forces that operate in the market are complex, and they affect the prices or shares of all the factors.

The prices of the factors of production are determined simultaneously. They are the resultant of all the forces operating in the market. If the price of one factor changes, this is bound to have some effect on the prices of all the others. To take a single example, suppose that wage rates rise because of a reduction in the supply of labor. This reduction might be due to an increase in the amount of disability among workers or to a very sharp decline in the birth rate. In any case, the higher wage rate would be sure to affect to some degree interest rates, rents, and profits. For example, it might make manufacturing less attractive to entrepreneurs, and so lower the interest rate by reducing the demand for capital. Again, it might increase the cost of operating farms and therefore cause a decrease in the demand for farm land, the abandonment of the poorest farms, and a fall in rents.

Some implications of distribution theory. A knowledge of distribution theory should aid us in understanding a number of important economic relationships. It is clear, for example, that, other things being equal, the price or share of a factor depends on its supply. The scarcer the factor, the higher its price and productivity per unit; the more plentiful the factor, the lower its price and productivity per unit. One important implication to be drawn from this relationship is that a heavy inflow of

immigrant workers into the United States would definitely tend to lower wage rates. This is not, however, as many assume, because immigrants are accustomed to low standards of living; rather because the inflow would increase the supply of labor.

In new countries it has often been observed that wage rates and interest rates tend to be relatively high. This is primarily because the supplies of labor and capital are scarce in relation to land. Because only the best land is used and comparatively little labor and capital are employed on each unit, the marginal productivity of labor and capital is high, and wages and interest rates are also high.

Finally, it should be noted that the more nearly the markets for the factors approach perfect competition, the less unemployment there is likely to be. In a perfectly competitive labor market, for example, everyone who wished to do so could sell his labor at the market rate and there would be no involuntary unemployment. The problem of unemployment in labor markets is the result of the many factors that interfere with effective competition. In extreme cases where a union has an actual monopoly of a certain kind of labor, the union is likely to set a relatively high price for the services of its members. The effect of such action, however, is to reduce employment below the level that would prevail had the wage rate not been raised. The reason for the decline in employment is that the individual employer will hire labor only up to the point where marginal productivity is equal to the wage rate; if wages are high, this point will be reached somewhat sooner than would be the case if they were lower.

How the national income of the United States is divided. By studying the financial statements of individual business concerns it is possible to note how in particular enterprises the product is divided between land, labor, capital, and entrepreneurship, although a precise allocation cannot always be made. It would be interesting to see how our national income is divided between the factors. Here again, however, we cannot have the data in the form that will give us the precise breakdown most desirable. However, the breakdown used by the Department of Commerce and shown in Table 27 gives us a good over-all view of distribution as it takes place in our national economy. Table 28 provides the same data on a percentage basis.

It is interesting to note the relative stability, percentagewise, of the various groups during this period of rapid increase in national income.

*Table 27.—National Income, 1939–1946, Received by Principal Groups * (Billions of dollars)*

	1939	1940	1941	1942	1943	1944	1945	1946
National income	72.5	81.3	103.8	136.5	168.3	182.3	182.8	178.2
Total compensation of employees †	47.8	51.8	64.3	84.7	109.1	121.2	122.9	116.8
Income of unincorporated enterprises ‡	11.3	12.6	16.5	22.7	26.0	27.7	30.2	34.9
Rental income of persons	3.5	3.6	4.3	5.4	6.1	6.7	6.9	6.9
Corporate profits §	5.7	9.2	14.6	19.8	23.7	23.5	19.7	16.4
Net interest	4.2	4.1	4.1	3.9	3.4	3.2	3.1	3.2

* SOURCE: Department of Commerce.

† Includes supplementary payments as well as salaries and wages.

‡ Includes farm as well as business and professional. Also includes inventory valuation adjustment.

§ Includes inventory valuation adjustment. Profits are before income tax.

Table 28.—Percentage of National Income Going to Principal Groups

	1939	1940	1941	1942	1943	1944	1945	1946
National income	100	100	100	100	100	100	100	100
Total compensation of employees	65.9	63.7	61.9	62.1	64.8	66.5	67.2	65.5
Income of unincorporated enterprises	15.6	15.5	15.9	16.6	15.5	15.2	16.5	19.6
Rental income of persons	4.8	4.4	4.1	4.0	3.6	3.7	3.8	3.9
Corporate profits	7.9	11.3	14.1	14.5	14.1	12.9	10.8	9.2
Net interest	5.8	5.1	4.0	2.8	2.0	1.7	1.7	1.8

Labor's percentage showed a tendency to increase after 1941 following a decline from 1932 to 1941. The percentage represented by corporate profits rose through 1942 and dropped back subsequently. Corporate profits actually retained were far less than those indicated in Table 27

and used in computing the percentages in Table 28 because of the high income-tax rates, particularly during the war years, that prevailed in this period. Other recipients of shares of the national income also paid income taxes.

Summary

In this chapter we have tried to answer the question, What determines the share of product that accrues to each of the factors of production? We found that the share of income going to each factor for its services is really a price, and that in a competitive market this price is determined, like any other price, by supply and demand.

It was noted, however, that there are special conditions that affect both the demand and the supply of the factors of production. The demand for a factor of production is derived from the demand for the product, or, more directly, from the value of the product. Only at the margin, however, can the products of the different factors be separated. The marginal productivity of a factor is the value of that part of the total output of an enterprise which depends on a single unit of the factor. If we assume that all units of the factor are alike, the schedule showing the marginal productivity of the factor when different amounts are employed is identical with the demand schedule for that factor. We may also represent the supply of a factor by a schedule. However, the conditions affecting supply are somewhat different for each of the factors. Once the demand and supply schedules for a factor are known, the equilibrium price can be determined. It will be at the point where the quantity demanded is equal to the quantity supplied. At this point price will equal marginal productivity. However, marginal productivity does not determine price. Rather, both are determined by all the market forces that influence either supply or demand.

Actually the units of a factor of production are seldom alike. Where they are different, the average price of the units and their average productivity at the margin are determined in the general market by the forces of supply and demand. However, the price of each individual unit will be determined by its marginal productivity.

The price of one factor of production cannot be determined independently of the prices of the other factors. A change in the price of any one factor will affect the prices of all the others. The prices of all the factors are determined simultaneously and are the resultant of all the forces that influence supply and demand.

In the following chapters we shall study in more detail the special problems involved in determining the price or share of each of the four major factors of production.

Study Questions

1. What is meant by the term distribution? Distinguish between distribution and marketing.
2. Into what basic shares are the payments to the factors of production divided?
3. What is the difference between functional and personal distribution?
4. Which kind of distribution is of primary concern to the student of economic principles? Why?
5. How are the prices of the factors of production determined?
6. What is meant by saying that productivity is the source of the demand for a factor?
7. What method would you use to determine what share is to go to labor, capital, land, or business enterprise in the production of an automobile, a chair, or a house?
8. What does marginal productivity have to do with the problem of distribution?
9. What is a marginal productivity schedule? Is there another name for this schedule?
10. How would you find the market demand and the market supply schedule of a factor of production?
11. Suppose various units of a certain factor, labor or land for example, are unlike. How would you find the prices of the unlike units?
12. What is meant by the statement that the economic share often differs from the contractual share? Under what conditions would this be true? Untrue?
13. Why are the shares accruing to labor, capital, land, and entrepreneurship all determined simultaneously?
14. Is the following statement correct? "The price or share of a factor depends upon its supply." Explain carefully.
15. Study Table 27. What conclusions can you draw from it? For example, if total compensation to employees increased, while shares to other factors remained constant or decreased, would this mean that labor was better off or that there were simply more laborers who received income?

Exercises

1. In 1944 the national income was divided roughly as follows: compensation to employees 66.5 per cent; income of unincorporated enterprises 15.2 per cent; interest 1.7 per cent; net corporate profits 12.9 per cent, and rental income of persons 3.7 per cent.

a. Is this functional or personal distribution?

b. Are these contractual or economic shares? Why or why not?

c. Do you think these shares are fair? Why or why not?

2. The *United States News* for Apr. 25, 1947, stated that as compared with 1939 prices, food prices were up 95 per cent; prices on house construction up 84 per cent; men's suits 72 per cent; automobiles 56 per cent. The article also pointed out that factory workers' yearly earnings were up 87 per cent.

a. What does this statement have to do with the problem of distribution? Explain carefully.

b. The *United States News* also pointed out that incomes of different groups varied widely. School teachers, for example, had less purchasing power than in 1939. Thus, although the average income in purchasing power was about $1,440 in 1939, in 1947 it was only $1,250. The automobile worker had $1,711 in purchasing power in 1939 and only $1,638 in 1947. Farmers were better off, having a 1947 net income up 47 per cent in terms of purchasing power over 1939. The construction worker had in 1947 purchasing power of $1,805 as compared with $1,570 in 1939. How would you reconcile these figures with the theory of distribution presented in the text?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 17.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Vol. I, Chap. 11.

Bye, R. T., and R. H. Blodgett: "Getting and Earning," F. S. Crofts & Co., New York, 1937.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 19.

Carver, R. N.: "The Distribution of Wealth," The Macmillan Company, New York, 1913, Chaps. 9, 12, and 13.

DEIBLER, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 17.

FAIRCHILD, F. R., E. S. FURNISS, and N. S. BUCK: "Economics," The Macmillan Company, New York, 1940, Chap. 15.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 14.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 23, pp. 379-381.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 25.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chap. 27.

21. Rent

The broad principles of distribution were outlined in the preceding chapter. The present chapter and the three following will explain how these principles apply to the share of each of the four major factors of production. The first share to be studied is *rent.*

Rent defined. As has already been noted, *rent is the share of income which represents payments made for the services of land.* This definition requires little amplification; it does, however, need a certain amount of explanation.

"Land" means natural resources; strictly speaking, it refers to gifts of nature in their original state, before labor has been applied to make them more useful to man. David Ricardo, an early nineteenth century English economist noted for his theory of rent, expressed this by defining 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." Ricardo was wrong, of course, in assuming that the "powers of the soil" are indestructible.[1] Nevertheless, most economists still use the term "land" to refer to the resources of the earth in their natural state, and we shall accept this common usage.

Very little land can be employed, however, unless improvements are made upon it. Some land must be cleared of forests or underbrush; other land must be drained or cleared of stones. Much labor and capital must be expended on urban sites before the land can be used for homes, stores, or factories. All these improvements on land constitute capital. Therefore part of the income from most pieces of land, for example from the typical farm, is really interest on invested capital; and in practice it is impossible to determine at all accurately how much of the income received should be considered interest and how much should be con-

[1] Doubtless what Ricardo meant was that rent is a payment based on the productivity of the land, on the assumption that the land itself is fully maintained and not used up. Ricardo would not have classified as rent a royalty on ore or oil taken from land.

sidered true rent. Nevertheless, a rough division can be made, and there are certain advantages in defining rent as income received for the use of the original powers of the earth: (1) This is the traditional sense in which economists have used the term. (2) The conditions that determine the supply of natural resources are somewhat different from those which determine the supply of capital goods produced with the aid of human labor.

It is important not to confuse true or economic rent with rent in the usual business sense. In everyday life we speak of renting a tractor, an office building, or a factory. Rent in this sense is a payment made for the use of any concrete unit of wealth over a definite period of time, regardless of whether the wealth is land or capital. True rent, however, consists only of payments made for the use of natural resources. This means that the "rent" paid for the use of a tractor is really not rent but *interest.* Of the "rent" paid for the use of an office building or a factory only the part paid for the use of the land on which the structure stands is true rent, while the part paid for the use of the building and equipment is really interest.[2]

Kinds of rent. There are several different rent concepts which it will be useful for the reader to understand and keep in mind.

First, he should note the distinction between *implicit* and *explicit* rent. A producer who uses his own land in the process of production seldom sets aside a special payment to himself in the form of rent, although his land makes a contribution to production and may contribute as much income as the owner could get if he leased it to someone else. We should recognize, therefore, that the factor land produces rent even when it is used by the owner. We reserve the term "implicit" rent for that part of the total product which must be allocated to land even when the land is used by its owner. If the landlord leases the land to another and receives a payment, this payment is called "explicit" rent. Obviously the same piece of land when used for similar purposes should yield the same rent, whether or not it is used by the owner. Thus implicit rent and explicit rent from a piece of land should be the same in all cases.

A second distinction is made between the contractual and economic shares that accrue to land. Since land is owned, the share that will be

[2] Not even all the payment made for the use of a building is *economic* or *pure* interest. This will be explained in a later chapter.

received by the owner is generally determined by contract. This is called *contractual rent.* On the other hand, the return that is just equivalent to the contribution of land to production is called *economic rent.* The contractual rent is the amount the owner actually receives for allowing someone to use his land; the economic rent is what he should receive. Under perfect competition and in the long run, these two shares tend to be equal. However, since both competition and human judgment are imperfect, the contractual share may often be out of line with the economic share.

Problem of explaining rent. In the past, economists have usually explained rent as a *surplus.* They have assumed that the prices of labor and capital were known, and also the prices of the goods produced. Then they have deducted labor and capital costs from the value of the total product, and explained that the "surplus" left over would go to land as rent. In this division of the product they seem largely to have ignored the entrepreneur. However, many economists have allowed him, as wages, the market value of his services, that is, the payment he could have obtained for them had he hired himself out to someone else.

The objection to this explanation of rent is that it assumes that the prices of labor and capital and the prices of the products of industry can be determined independently of rent. Actually this is impossible. Wage rates, interest rates, and rents are all interdependent; they influence one another and, in the final analysis, are all controlled by the same basic economic forces. To illustrate this fact it is only necessary to call attention to the nature of the demand for the factors of production. The demand for the factors is derived from the demand for their products; but since no production whatever is possible unless all the factors are employed in combination, the demand for the factors is a joint demand. If the demand for a product of land, labor, and capital increases, with a resultant increase in the price of the product, this will tend to raise not only rents but also wages and interest. Suppose, for example, that a war brings a great increase in the demand for wheat and a sharp rise in its price. This will increase rents by increasing the demand for land; and it will also increase wages and interest rates in the wheat-growing areas by increasing the demand for labor and capital.

Basically, therefore, the theory of rent determination is no different from the theory of the determination of wages and interest. Like wages and interest, rent is a price, and in a perfectly competitive market it

would be determined by supply and demand at the point where the rental price was equal to the marginal productivity of land.

Undoubtedly one reason that so many writers of a generation or two ago failed to see that the rent problem was essentially no different from the problem of wages or the problem of interest was this: In their first approach to the determination of wages and interest they usually thought of labor and capital as uniform commodities for each of which a single price could be found. In other words, they were trying to explain *the* wage rate and *the* interest rate. In the actual world, however, there are a number of wage rates and a number of interest rates; therefore what they really explained, insofar as their theories were sound, was the determination of the *general level* of wages or interest. On the other hand, when they approached the problem of rent determination, they were thinking first of all in terms of *differences* in the quality of land and *differences* in rent; as a result, their rent theory was chiefly concerned, not with explaining the general level of rents, but with explaining why one piece of land rented for more than another.

Once the general level of wages, interest, rents, and marginal products, as well as the prices of the final goods, has been brought into equilibrium by the forces of the market, it is quite legitimate to explain the price of *any one unit* of a factor of production by its marginal productivity. We shall see later that one way to find the marginal productivity of a particular unit of land is to subtract labor and capital costs from the value of the total product.

In our explanation of rent determination in this chapter we shall be concerned with two quite distinct problems. First, we shall explain how the general level of rents is determined by the forces of supply and demand. In order to simplify as much as possible our explanation of the general level of rents, we shall assume that land is a uniform commodity, so that every unit is exactly like every other unit. Of course land is very far from being a uniform commodity. Nevertheless, it is reasonable to assume that the general level of land rents responds to the forces of supply and demand much as would our imaginary unit based on the assumption of uniformity. The second problem that will concern us is to explain why rents vary on different pieces of land. This is an important problem, because differences in rent are great. In this portion of our discussion we shall use the traditional type of rent analysis.

Demand for land. In explaining the general level of rents let us assume that our unit of land is an acre of farm land, and that all units are exactly alike in fertility and possess equally favorable locations. Further, let us assume that the quantities of the other factors, and the prices of the products of the factors, remain constant. We can then represent a demand schedule for land exactly as we represented a demand schedule for labor. Each point on the demand curve will show the number of acres that will be demanded at a given price; each point on the demand curve will also show the marginal productivity of an acre when a certain number of acres are employed in the economy as a whole. In other words the demand curve, as in the case of the demand curve for labor, will also be a marginal productivity curve.

Supply of land. It is when we come to supply that we find some important differences between land and the other factors. The stock of capital in the world, and the labor force, may change with the passage of time; but the "stock" of land is essentially fixed. There are, to be sure, some rather important qualifications to be made to this statement. Land in the economic sense is not just the surface of the earth, but is natural resources. Resources are things that are useful to man, and there are several ways in which the stock of natural resources may be changed. In the first place, natural resources are not indestructible, but can often be used up, or destroyed by misuse. In the second place, natural resources may exist of which we have no knowledge, especially in the case of mineral deposits. When these are unexpectedly discovered, our total stock of natural resources is increased just as effectively as if we had had a gift from Mars. In the third place, materials that are not natural resources are constantly being brought into that category by the advances of science and technology. Before the development of kerosene and the kerosene lamp, petroleum was not a natural resource; neither was bauxite, the ore from which aluminum is obtained, until methods were discovered for extracting the metal from the ore. Nevertheless it remains true that our most basic natural resource, the land area of the world from which we can obtain farm and forest products, is very definitely limited. Any possible additions to it are of relatively minor importance.

The supply of a factor of production is not, to be sure, the known quantities in existence; rather it is a schedule of the quantities that would be offered in the market at various prices. However, since the available

quantity of land is essentially fixed, irrespective of its rental price, it is reasonable to suppose that in the long run most landowners would rent their land for any price they could get rather than let it remain idle. Therefore, we should expect the supply schedule for the services of land to be highly inelastic. It follows, therefore, that at any price practically the whole quantity available would be offered in the rental market.

Determination of the rental price. Under the assumption that all units of land are alike, the rent of an acre of land will be determined by supply

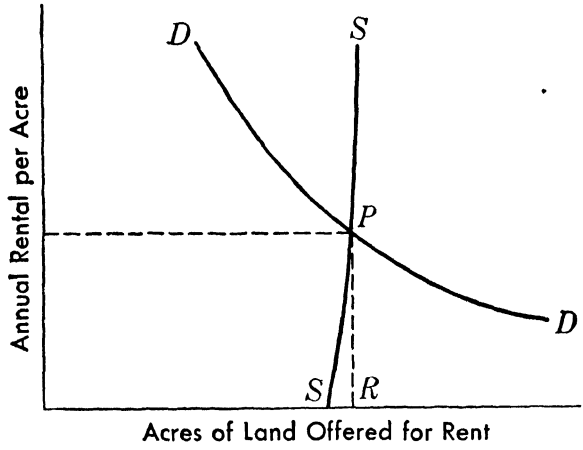


Fig. 1.—Determination of rent when all land consists of identical acre units.

and demand at the point where the number of acres offered is equal to the number of acres demanded. At this point the price of an acre will be equal to its marginal productivity.

This situation is represented in Fig. 1. This chart shows that the annual rental of an acre of land under the conditions represented would be *PR*. *PR* also represents the marginal productivity of land. Since the supply of land is practically fixed, over a period of time demand will be the active force in determining rent. If population and capital increase, the demand curve *DD* will gradually shift to the right, and rent will rise. Though the chart is based on the assumption that all units of land are alike, it gives us some concept of how supply and demand determine the *average* level of productivity of land at the margin, and also the *average* level of rents even when land is not uniform. Lack of uniformity in the units of a commodity does not seriously interfere with the opera-

tion of the forces of supply and demand. It merely prevents the establishment of a single price.

Rent on unlike units of land. As everyone knows, units of land differ widely in quality and productivity, and in the amount of rent that they will yield. It is therefore a matter of some importance that the student of economics should understand how rent is determined on particular pieces of land. It has already been noted that once interest rates, wage rates, the general level of rents, and the prices of products have all been brought into equilibrium by the forces of supply and demand, it is then possible to find the marginal productivity and the rent of a particular unit of land by subtracting capital and labor costs from the total value of the product. The reasons that this is true will be made clear later. Subtracting labor and capital costs from the value of the product is the traditional procedure for analyzing rent. Though it does not explain how the general level of prices, costs, and rents is determined, it does give a clear idea of why rent arises on some pieces of land and not on others, and of how the amount of rent on each piece of land is determined.

Before we can discover the marginal productivity and the rent of land by subtracting labor and capital costs from the value of the product, we must first discover how much labor and capital it will pay to employ on a piece of land and what the total value of the product will then be. To simplify the exposition we shall assume that labor and capital must be applied in combinations or units containing a certain amount of each. This will enable us to treat labor and capital as if they were a single factor of production. We shall then be dealing with only two factors: labor and capital on the one hand and land on the other. To simplify the exposition further, we shall assume that all land is farm land and that the unit of land is the acre.

There are two conceivable ways in which a given unit of labor and capital can be employed to produce an output. Either it can be put to work with other units of labor and capital on an acre of land which is already being cultivated, or it can be put to work on an acre of land, if one is available, which has not yet been brought into use. Increasing output by applying labor and capital to larger areas of land is called *extensive cultivation.* Increasing output by applying more labor and capital to the same land is called *intensive cultivation.*

In explaining how the rent of individual units of land is determined, we shall find it useful, first, to assume that only one unit of labor and

capital can be employed on each unit of land. This will mean, of course, that a unit of labor and capital can be made productive only by the method of extensive cultivation. Later we shall show how the results obtained must be modified by taking account of intensive cultivation. Intensive cultivation is no less important than extensive cultivation, because the rent on a good piece of land is explained, not only by the fact that *one* unit of labor and capital can produce more on the good land than on the poor land, but also by the fact that on the good land it pays to use additional units of labor and capital. In due course our analysis will show why this is true.

Rents on different units of land under extensive cultivation. Let us assume that the various acres of land are not equal in fertility; that is, an application of a unit of labor and capital on each plot will produce different amounts of output. Let us also assume that only one unit of labor and capital can be applied to any 1 acre of land. To increase production it will then be necessary to employ extensive cultivation and bring additional acres of land into use. Some of these assumptions are, of course, artificial; yet by their use we can demonstrate more simply the reasons that, under conditions of extensive cultivation, different units of land produce different amounts of rent.

In view of the assumption of unequal fertility, labor and capital applied to some pieces of land will be more productive than when applied to others. It is reasonable to suppose, therefore, that producers will be willing to pay owners of the more fertile land for its use even though less fertile land may be available without cost.

In order to demonstrate that rent will be paid on the better pieces of land under conditions of extensive cultivation, and also that it will be different on different units, the following illustration is presented. Assume that six units of land of unequal fertility are not now in use but are available for raising wheat. Assume that the probable results of employing one unit of labor and capital on each plot will be as follows: land A will yield a total of 50 bushels; B, 40; C, 30; D, 20; E, 10; and F, 5. Suppose that an acre of land will be brought under cultivation by someone provided it will produce at least enough wheat to cover labor and capital costs. Suppose further that the price of wheat as determined in the market is $1 per bushel, and that the price of a unit of labor and capital is $10. In these circumstances it clearly will not pay to use land F at all. Land E, however, will be brought under cultivation, since the

price of the product will just cover labor and capital costs. However, since even on land E no return will be secured over and above capital costs, no one will be willing to pay for the use of land E and it will receive no rent. The situation is different for land units A, B, C, and D. If the price of wheat is $1 per bushel, the user of land A will need to invest only $10 in labor and capital to secure a gross return of $50. This will mean an excess of $40 above labor and capital costs. If necessary, in order to have the privilege of cultivating land A, the prospective user will surely agree to pay part of the $40 "surplus" to the owner of the land as rent. Similarly a user of land B would receive a gross income of $30 in excess of labor and capital costs; of land C, $20; and of land D, $10. Each of these prospective users would therefore be willing, if necessary, to pay at least a part of this excess income to the landowner.

The users of superior pieces of land will not, of course, pay rent unless they have to. How much will they actually be obliged to pay? Will the user of land A have to pay as rent the entire $40 by which the value of the product exceeds labor and capital costs? If he pays the whole $40, he might just as well be cultivating land E, for he will secure a return on his labor and capital only. The other prospective renters will be in a similar position if the rent on B land is $30, on C $20, and on D $10. Yet if we assume a free competitive market and perfect mobility of the factors of production, the users of all these plots of land will have to pay as rent the full amount by which the value of the product exceeds labor and capital costs; for no one who has labor and capital at his command can hope to receive more for their services than the prevailing rate. The various producers who wish to put their labor and capital to work will engage in competitive bidding for the use of land. If in the beginning the rent of a piece of land were low enough so that a user could make a net economic profit after receiving wages and interest on his capital at the market rate, some prospective tenants would be willing to offer a little more rent in order to get the use of such land. As a result, competition for superior land would force those who desired to use it to pay the owners of the land the full amount by which the value of the product exceeded labor and capital costs. Thus rent would be paid on land units A, B, C, and D. Moreover, a different amount of rent would be paid for each unit according to its superiority in productive power.

Figure 2 shows the different rents that result when plots of land of different fertility are cultivated extensively. It illustrates graphically how

differences in rent arise. We see that, under extensive cultivation only, rents will vary in amount on different pieces of land because the application of the same amounts of labor and capital to different plots will yield different results.

Marginal land. We observe that E land is cultivated but that the return is sufficient only to pay for the labor and capital costs involved. Hence, there is no marginal product attributable to land E and no income for the owner. When returns to producers who use land are sufficient to

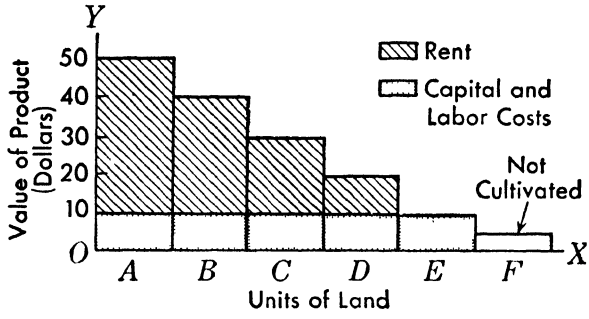


Fig. 2.—Rents on land of unequal fertility on the assumption that only extensive cultivation is possible.

pay only for labor and capital costs, the land is called *marginal* or *no-rent* land. If land is so poor that it will not even pay labor and capital costs like the F land in our illustration, it is called *submarginal* land. Such land will not be used at all as long as prices do not increase.

It must, however, not be assumed that the units of land that are marginal at one time necessarily remain so. If market conditions change in such a way as to change the prices of the factors or the prices of their products, the units of land found at the margin will also change. For example, if the price of the product should rise without a corresponding rise in wages and interest, it would pay to cultivate somewhat poorer grades of land, and land units that had formerly been submarginal would become marginal. Also, units that had formerly been on the margin would begin to yield rent. Again, if the costs of labor and capital should rise without a rise in the price of the product, land that had been marginal would no longer cover labor and capital costs and would be abandoned; and land units that had previously been above the margin would then become the marginal units.

Differences in rents when cultivation is intensive as well as extensive. So far we have studied differences in rent on the assumption that the same amount of labor and capital, namely, one unit, is applied on every acre of land. Actually, however, if it pays to use the poorer plots of land at all, it pays to use greater amounts of capital and labor on the superior plots; and it is this intensive cultivation of the better units of land that accounts in large part for their high marginal productivity and high rents.

In our illustration of how the rents of individual units of land will be determined if only extensive cultivation is possible, we assumed that the forces of supply and demand in the market had determined the price of a unit of capital and labor at $10 and the price of a bushel of wheat at $1. We then found that one of these $10 units of capital and labor would produce $50 worth of wheat on plot A, which represented land of the highest grade. That being true, and even assuming that the law of diminishing returns begins to operate after the application of the first unit of labor and capital, it is pretty certain that a second unit of labor and capital applied on plot A would increase the value of the product by considerably more than the $10 that the unit would cost. For purposes of illustration we shall assume that a second unit would add $40 to the value of the product, a third unit $30, a fourth $20, and a fifth $10.

Table 29 shows the additions that would be made to the value of the product by successive $10 investments of labor and capital, not only on land A, but also on each of the other plots of land. In addition, assuming that labor and capital would be added to each plot up to the point where the additional product just covered the additional cost, Table 29 shows for each plot the total value of the product, the total labor and capital cost, and the rent. The heavy broken line in the table shows how many units of capital and labor should be employed on each plot; and it also divides the potential products that would be produced from those which would not.

It is clear from the table that on plot A five units of labor and capital will be employed. Whether the fifth unit is employed is really a matter of indifference, but we have assumed that any unit will be employed if the value of the added product just covers the additional cost.[3] When

[3] According to mathematical theory, if capital and labor could be divided into *infinitesimal* units, it would pay to use more capital and labor *until* the marginal cost was just equal to the marginal productivity.

five units of labor and capital are employed, the total value of the product will be $150, and the total labor and capital cost $50. The rent on A land will therefore be $100, since rent can be measured by the difference between the total value of the product and the labor and capital cost.

Table 29.—Rent under Intensive and Extensive Cultivation

Plot	Units of Labor and Capital Applied					Total Value of Product	Total Cost of Labor and Capital	Rent
	1st	2nd	3rd	4th	5th			
A	50	40	30	20	10	$150	$50	$100
B	40	30	20	10	8	100	40	60
C	30	20	10	8	6	60	30	30
D	20	10	8	6	4	30	20	10
E	10	8	6	4	2	10	10	0
F	5	4	3	2	1	0	0	0

In a similar fashion the rent on each of the other acre plots of land can be calculated. E land will be cultivated but will yield no rent, and therefore it is marginal. F land is submarginal, for it cannot be cultivated at all except at a loss.

It should be noted that the productivity of labor and capital tends to be the same at both the intensive and the extensive margins; that is, the one unit employed on the marginal plot of land E yields the same value of product ($10) as the last unit employed on each of the other plots of land.

Formula for finding the rent of a unit of land. In Table 29 the rent of each plot of land was found by subtracting labor and capital costs from the total value of the product, that is to say, from total income. This can be expressed as a very simple formula. Let R stand for rent, T for total income, and C for the cost of labor and capital. The formula then becomes $R = T - C$. Of course, to apply this formula labor and capital must have been employed in the most advantageous quantities, until the marginal cost of the labor and capital was just equal to the value of the marginal product.

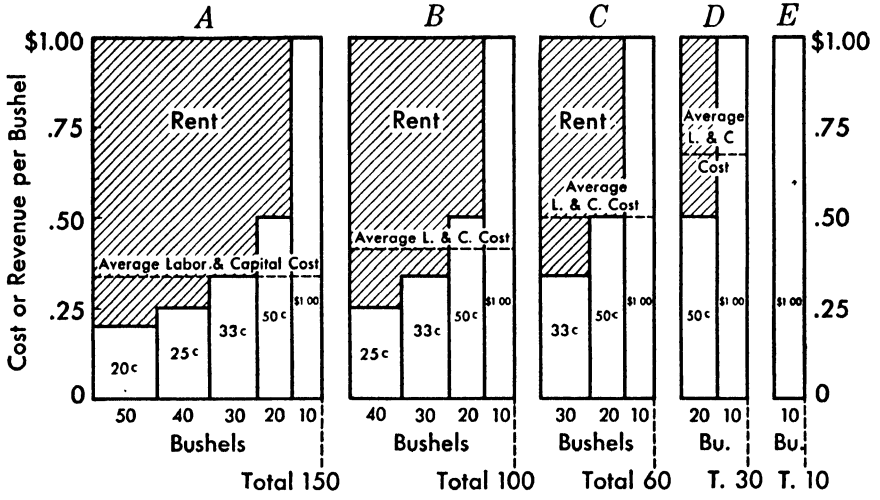


Fig. 3.—Relationship between average labor and capital costs, marginal labor and capital costs, bushels of wheat produced, and rent.

Relation of average cost to marginal cost and rent. A chart showing the relationships between average labor and capital costs, marginal labor and capital costs, and rent may give the reader a clearer understanding of how rents are determined on individual pieces of land. Figure 3 shows these relationships for all the plots of land in Table 29 except F. Plot F is not included because it does not pay to produce anything on it. The base line of each column shows the increment of wheat attributed to each successive unit of labor and capital. The heights of the columns show the average labor and capital costs of each increment. The broken horizontal lines cutting the columns show the average labor and capital costs of wheat on each plot when production has been pushed to the

limit. The shaded areas represent rent. Rent could also, alternatively, be represented by the areas above the broken average cost lines.

The chart shows how each producer will expand output until the labor and capital cost of the last unit of product just equals the price in the market. It is, of course, marginal and not average costs that are significant in determining how intensively a piece of land will be cultivated. On the better pieces of land average cost will always be below marginal cost. Thus, on A land average labor and capital costs are only 33⅓ cents when production reaches its limit; on B land they are only 40 cents, and so on. However, marginal labor and capital costs are always equal to the selling price.

Rents on residential and commercial land. Part of the payment made for using a house or a store is rent for the land on which it is built. The general principles of supply and demand apply to rents on residential and commercial sites as well as to farm land. Residence sites, however, are a consumers' good. Therefore the demand for their services is not derived from the demand for a product; rather, their services are demanded because they directly satisfy consumers' wants. Differences in rent on residence sites must, then, be explained in terms of differences in desirability and not in terms of differences in productivity.

Commercial sites are more like farm land in that the demand for them is derived from their productivity, that is, their power to yield income. Just as a good acre of farm land will yield more income than a poor acre because it will produce more wheat, so a good drugstore site will yield more income than a poor one because it will produce more sales. In both cases a single unit of labor and capital will be more productive on the "good" land than on the "poor" land; and in both cases it will pay to use more units of capital and labor on the superior plot.

Effect of rent on prices. In former times writers on economics often maintained that rent was not a "price-determining" cost. No modern economist, however, would subscribe to this doctrine without very important qualifications. In the first place, costs do not *determine* prices; that is, not in the sense that costs are something independent and fixed to which prices must become adjusted. Costs themselves are prices—the prices of the factors of production. As such, they are influenced by the demand for and the prices of the finished products. For example, if an increase in demand raises the prices of textile products, this will tend to

increase the demand for raw cotton and for labor in the textile mills, and will raise their prices also.

Nevertheless, it is true that, if an entrepreneur wishes to produce a given commodity, like clothespins, he must pay the market prices for labor and wood; and the clothespin industry has very little effect on these prices. Therefore, since the price of a commodity tends in the long run to equal average costs of production, it is essentially correct to say that the price of clothespins will be determined by cost. But if this is what we mean when we say that cost of production determines price, rent is price-determining as much as any other cost. To get a desirable factory site the clothespin manufacturer must pay at least as much rent as others competing for the land—say farmers or merchants—are willing to offer. This is the "opportunity cost" of the land. In the long run the price of clothespins must cover rent as well as other costs or clothespins cannot be produced.

Those who maintain that rent is not a price-determining cost are really assuming that land produces only a single product. They argue that the demand for land is derived from the demand for, and the price of, the product. If the price of the product is low and the demand for land very small, land will be a free good. As the price of the product rises, the demand for land will increase and rent will appear on the superior units. Therefore, they maintain, rent is a *price-determined,* not a *price-determining,* cost.

The weakness of this argument is that practically all land has a number of alternative uses, and therefore rent appears on desirable units whether any particular product is produced or not. For example, there is little if any land suitable for raising wheat or cotton that is not also suitable for the production of various other farm products. Many products raised on farms require relatively good land and cannot be produced at all unless their prices are sufficiently high to cover rents that the land will yield if put to other uses. Few if any products can be produced in any quantity on land that is unsuitable for all other purposes, or that is marginal for all other purposes.

Capital value of land. The capital value of land used in production is derived from the income that it yields, that is to say, from its rent. Under certain conditions it is possible to estimate the capital value of a unit of land rather accurately on the basis of the rental yield. To do this it is necessary (1) to have such stable economic conditions that the rent

can be expected to remain constant for a very long time; (2) to know the rate of interest on investments yielding an assured income and to know that this rate of interest will also remain constant.

Let us assume that the current rate of interest on investments that will yield an assured and stable income is 3 per cent, and that a certain piece of land can be expected indefinitely to yield a net rental income of $1,200 per year. With this information we can find the market value of the land by capitalizing the rental income of $1,200 at the interest rate of 3 per cent. In a competitive market the land will tend to sell at such a price that $1,200 will equal 3 per cent of the capital value. If $1,200 is 3 per cent of this value, the capital value itself will be $40,000. We get this result by dividing $1,200 by 0.03. This means that *the value of land equals the rent divided by the rate of interest.* If we let V be the value of the land, R the rent, and I the rate of interest, we can express the value of land by the simple formula $V = R/I$. Applied to our illustration this becomes $V = \$1{,}200/0.03 = \$40{,}000$.

In the real world, of course, there is no assurance that either land rents or interest rates will remain constant, and in estimating a fair market value for land many special circumstances have to be taken into account. Nevertheless, capitalization of the rental income is a useful guide for those who must estimate land values.

Summary

Rent is the price paid for the services of land, and, like wages and interest, is determined in the market by the forces of supply and demand. If we assume that land is a uniform commodity, the demand for it can be expressed as a schedule of the quantities that will be employed at various rental prices. Since the demand for land is derived from its productivity at the margin, a demand schedule for land is identical with a marginal productivity schedule for land. The supply of land can also be represented by a schedule, but this schedule will be highly inelastic and subject to little change because the total "stock" of land is pretty definitely fixed by nature.

Interest, wages, the general level of rents, and the prices of products are all determined simultaneously by the forces of supply and demand in the market. However, once interest rates, wage rates, and the prices of products have been thus determined, the differences in rents on units of land of different quality can be explained in terms of differences in marginal productivity. Since variations in the quality of land are great, it is important to understand

how the rent of particular units of land is determined. If labor and capital have been employed up to the point where their marginal cost is equal to their marginal productivity, the rent of any given unit of land can be found by subtracting labor and capital costs from the total value of the product. What remains is the marginal productivity of the land, that is to say, its rent.

Among the facts about rent that the reader should keep in mind are the following: (1) Rents on land used for residential and commercial purposes can be explained on much the same principles as rents on farm land. (2) Contrary to the belief of some economists of a generation or two ago, rents have much the same effect on prices as do other costs of production. (3) If the rent on a piece of land is known and the interest rate is known, and both are stable, it is possible to estimate the capital value of the land by capitalizing the rent at the interest rate.

Study Questions

1. Define the term rent.
2. What is meant when the term land is used?
3. Who was David Ricardo and how did he define the term rent? Criticize his definition.
4. In what sense and to what extent is the stock of land or natural resources fixed? Explain carefully.
5. What is the distinction between rent as used by the economist and the term when used in the business sense? Which definition is being used in this chapter?
6. Distinguish between implicit rent and explicit rent. Is the economist in discussing principles concerned with either or both of these concepts?
7. Distinguish between contractual rent and economic rent.
8. What problems arise when attempts are made to explain rent? Why do these problems arise?
9. In what sense is rent a surplus? Explain carefully.
10. What is meant by saying that the demand for the factors of production is a joint demand? How does this affect the explanation of rent?
11. What two distinct problems relative to the determination of rent are considered in this chapter? Why is it necessary to consider these problems?
12. What is meant by the general level of rents?

13. How is the demand for land determined when a general level of rents is to be found?

14. Does the supply of land differ in any respect from the supply of capital or labor? Explain.

15. Why is it true that demand is the active force in determining rent?

16. How would you determine the rent of unlike units of land? What assumptions are made in this explanation?

17. Distinguish between the intensive and extensive margins of cultivation.

18. What is meant by the statement that the productivity of labor and of capital tend to be equal both at the intensive and at the extensive margins?

19. Will no-rent land ever be used? Explain. Would you cultivate no-rent land extensively or intensively, or both? Explain.

20. What is marginal land? submarginal land?

21. State the formula for finding the rent of any one unit of land. Apply the formula to land B in Table 29. Is your answer similar to that shown in the table?

22. Why are marginal and not average costs significant in determining how intensively a piece of land will be cultivated?

23. Is the principle of rent determination any different for residential or commercial land from that for agricultural land? Explain.

24. Is rent price determining? Explain carefully.

25. How does the principle of opportunity costs apply to the problem of rent?

26. Suppose a piece of land yields a net average return of $200 per year. Assume an interest rate of 5 per cent. How much should be paid for this land? Why? Suppose the rate of interest falls to 3 per cent. Will the land be more or less valuable? Why?

Exercises

1. Assume that an entrepreneur had his choice of five pieces of land upon which to produce a cereal crop. Land A would yield 50 bushels; land B, 45; land C, 40; land D, 35; and land E, 30. The cost of all other factors (except land) in producing the crop is equivalent to 35 bushels.

a. Which of the plots would it be worth while to cultivate?

b. Which is the marginal unit of land?

c. How much rent could the owner of each plot claim from the entrepreneur?

d. Assuming that the real rent was claimed in each case, which plot of land would the entrepreneur wish to use if he could cultivate only one?

2. Five pieces of land are available for the production of a certain commodity. The land is cultivated with the following results. Notice both extensive and intensive cultivation.

Units of labor and capital applied	*Product per unit applied*				
	A	B	C	D	E
First	50	45	40	35	30
Second	45	40	35	30	25
Third	40	35	30	25	20
Fourth	35	30	25	20	15
Fifth	30	25	20	15	10
Sixth	25	20	15	10	5
Seventh	20	15	10	5	0

a. Labor and capital costs $30 per unit of application, while the product sells for $1 per unit. How many units of labor and capital will be applied to each piece of land?

b. Suppose the price of the product rises to $2 per unit while the cost of each application of labor and capital remains at $30. How many units of labor and capital will be applied to each piece of land so far as you can determine it from the table?

c. If only six units of labor and capital are available but all of the land can be used, how would you allocate the units of labor and capital to the different pieces of land? State precisely the principle that you follow in making the allocation. Does this illustrate what is meant by the statement that at the intensive as well as at the extensive margin of cultivation production is equal?

d. Under cost and price conditions in (*a*), how much rent would you pay for each of the pieces of land? Show by a table how you obtained your results.

3. Given the following data:

Number of units of labor and capital employed	7
Total yield of product	80
Yield at the intensive margin	8
Value of product per unit	$2

Compute:

a. Marginal product of labor and capital.

b. Value of a unit of labor and capital.

c. Rent of this piece of land.

d. Suppose that the interest rate is 5 per cent, how much would you pay for this piece of land?

4. Mr. Jones has a farm that normally yields 25 bushels of wheat per acre. The product normally sells for $1 per bushel. Suppose Jones can cultivate 100 acres of land. With the same labor and capital he could sell his services in other pursuits for $1,800 per year.

a. What is the marginal cost of wheat in the country per bushel?

b. How much is this 100-acre farm worth if the interest rate is 5 per cent?

c. Suppose that there is a temporary corner on wheat so that the price for a week is $2 per bushel. How would this increase in price affect the value of this farm?

d. Suppose that Jones discovers that, because of an increase in demand for labor and capital in other pursuits, he could earn elsewhere $3,000 per year. Suppose that the price of wheat remains at $1 per bushel. Should he continue to farm?

e. If the conditions in (*d*) are permanent, what would this farm be worth as a productive unit?

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 18.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 22.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chaps. 18 and 19.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chap. 15.

FROMAN, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 15.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 25.

GEMMILL, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 26.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chap. 30.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chap. 28.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. II, Chaps. 44-46.

23. Interest

The return received by capital, another factor of production, is known as interest, and in this chapter we shall learn how the rate of interest is determined. First, however, it is necessary to understand what is meant by the term "capital." Capital was defined in an earlier chapter as all produced goods that are used in further production. Strictly speaking, capital is not money. However, for many purposes it is convenient to count as capital the liquid funds employed in carrying on production. In this chapter, therefore, we shall treat such funds as one form of capital. Capital thus includes capital instruments, goods in process, finished goods awaiting sale, and the funds that are required in the day-to-day operation of a business. Since capital is produced, it is distinguished from land, which is a natural resource. It is also distinguished from consumers' goods, because these are consumed directly instead of being used in further production.

Capital should not be confused with wealth. Wealth is defined as a stock of economic goods existing at a particular time. Wealth, therefore, includes not only capital but also land and consumers' goods. A man, for example, who owns a factory, a piece of land, and a stock of food and clothing for his own use possesses wealth, but only the factory and machinery are capital. Likewise capital and money do not always mean the same thing. Capital goods and capital funds are employed in production. Money, on the other hand, is merely a medium of exchange. It is also the standard in which we express the value of capital goods and other forms of wealth. Money is a form of capital only when it is used in the production or distribution of goods.

Some individuals, of course, have only *money* to invest. Such money is called *liquid* or *free* capital. It is capital funds as contrasted with capital goods. Those who have money to invest may either purchase capital goods or allow someone else to invest their funds. In either case, the money is finally invested in capital goods which furnish the income paid

to the lenders. The owner of capital funds, for example, may not want to build a new pipe line for gas from Texas to New York, but he would be willing to lend funds to a company engaged in the transmission of natural gas.

Interest defined. Just as the term "rent" is defined as the net annual return for the services of land, so interest may be defined as the *net annual return to capital, expressed in terms of a rate or a percentage.* The actual payment received by the owner of capital, however, often includes something more than compensation for the services of capital goods. If an owner lends a capital instrument, or its equivalent in money, the total payment he receives includes many items. The payment should, for example, cover the wear and tear on the instrument, taxes, and insurance. Also it should compensate the owner for certain uninsurable risks and for his trouble in making the loan and supervising it. All these items are cost to the capitalist, but they often are included in the return that is receivable from capital. However, these costs must be subtracted from the annual income before the net or pure interest can be determined.

It should be noted that the income that the owner of a capital instrument receives is generally in the form of money paid to him periodically as a lump sum. Yet interest actually is a rate of return, or a percentage of capital. Before the percentage can be determined in any one situation, we must know the monetary value of a capital good. This can be found either by using the original cost of the good as its present value or by making a new evaluation at current prices. Thus, if the price level has increased, we may consider that the money value of a capital instrument has risen also; if prices have fallen, the money value of the capital good has decreased. The rate of return can be found by dividing the pure interest received each year by the value of the capital good.

Another method often used to estimate the money value of a capital good is to capitalize the annual income received. The use of this method, however, leads to circular reasoning when the rate of interest received from capital is sought. It is clear that a rate cannot be found unless the value of a capital instrument is known; but finding the value of the instrument by capitalizing the income requires knowledge of the rate. This method can be used, however, when a certain rate of return has been adopted as representing the current rate of interest. For example, the rate of return on long-term government bonds sometimes is regarded as an indicator of the pure rate of interest prevailing at the moment. If

some such rate is accepted, the value of a capital instrument can be found by dividing the annual income, less expenses, by the rate of interest. It should be noted, however, that the value of a capital good cannot for long depart very much from the cost of producing others like it. No one will pay a large amount for a capital instrument if he can secure an exact duplicate for a smaller amount of money; hence the market value of a capital good will tend, in the long run at least, to correspond fairly closely to its reproduction cost.

One further point should be noted. In this discussion our primary concern is to determine the net return for the services of capital used in production, whether in the form of goods or funds. Not all loans, however, are made for production. Many individuals borrow money or goods for consumption. For example, a man may borrow money to purchase an automobile for pleasure. He, too, must pay a return on the money that he has borrowed, and this return is called interest. However, the amount of consumption loans, when compared with those made for production, is so small that the rate paid on them depends largely on that paid for productive loans. Therefore, we need not take loans for consumption into account in our study of the return received by capital.

Contractual interest versus the pure rate of interest. As we have noted, owners of capital cannot regard as interest all the return that they receive. Their total income includes compensation for taxes, risk, inconvenience, insurance, and management in addition to a return on capital. We use the term "pure interest" to mean the payment to capital for its services after all other costs have been subtracted. *Contractual interest,* however, is the total payment made by borrower to lender. It includes not only the pure interest but all costs as well.

The distinction between pure interest and contractual interest explains why many different interest rates are found in the market at the same time. Contractual rates, for example, can vary because of differences in costs of supervision. On the other hand, some differences in interest rates can be explained by the length of time for which money is lent. Other differences arise because of taxes. In the main, however, contractual interest rates vary with differences in risk. Although the pure interest rate may be 3 per cent, the investor may demand 8 or 10 per cent if the risk is great. This extra amount, although often called interest, really is an insurance premium against loss of capital; that is, the owner of capital demands a high rate of interest for fear he may lose his principal.

Variation in contractual interest rates does not affect the principles that determine the pure interest rate. It is important, however, that the differences between contractual rates and the pure interest rate be recognized, for failure to see this difference leads to confusion. It is sufficient for our purposes to understand that the pure interest rate contains only a return on capital and that the contractual rate may contain many elements. The pure rate of interest is the rate that could be obtained on a permanent investment involving no risk, trouble, or expense.

Interest rate. We can say that the rate of interest on capital in a certain country in a certain year is determined by demand and supply. The demand, for the most part, comes from businessmen who wish to borrow capital for productive purposes. The supply is provided by owners of funds. Individuals and business concerns have accumulated capital through saving in the past, or they may be accumulating capital now by saving. Money made available through operations of commercial banks is also an important source of supply.

Under competition, the equilibrium rate of interest will be that rate which equalizes supply and demand. At that rate, the amount wanted by borrowers will equal the amount offered by suppliers.

As was true in connection with rent and wages, interest cannot be determined independently of the return paid to the other factors. The businessman who borrows capital also hires labor and rents natural resources such as land. He uses all the factors in making a product that he expects to sell. From the proceeds of the sale he reimburses himself for the compensation paid to the other factors and for his own efforts. Thus the most that he can afford to pay any factor is that factor's contribution to production in combination with the other factors. The returns to capital, land, labor, and entrepreneurship are therefore determined simultaneously, not independently of each other.

Since the rate of interest is governed by both demand and supply, we must consider further the reasons why businessmen elect to borrow and the reasons why suppliers are willing to accumulate capital. Looking at the demand side, we may ask the question, Why can interest be paid? From the supply side, we shall have to answer the question, Why must interest be paid?

Demand for capital. Why can interest be paid? The obvious answer to this inquiry is that capital is productive. This statement, however, requires explanation. For centuries, thinkers have questioned the pro-

ductivity of capital funds, or money. Aristotle maintained that "money is barren," and his conclusion was accepted for many years. The Church during the Middle Ages condemned the taking of interest because interest was supposed to be a payment based on fraud or force. Even today many persons believe that interest should not be paid. Our usury laws testify to the feeling against the taking of interest at high rates. Socialists and labor leaders sometimes consider interest as one form of the exploitation of labor.

In this connection two points should be noted: (1) Those who consider the taking of interest as a fraud or exploitation often confuse contractual with economic interest. Ample historical evidence indicates that lenders often have been in a monopolistic position and could secure an excessively high contractual rate of return. (2) Confusion arises because interest is supposed to be received for the use of money. Hence, since the state in a degree controls the issuance of money, some persons assume that all demands for loans could be satisfied without payment of interest if the state would create enough money. This argument will be considered later in the chapter.

If we accept the theory that capital is productive, we can explain why interest is paid. Interest is paid because capital instruments are in demand. Why are capital instruments demanded? The answer is that capital instruments, when used with land, labor, and entrepreneurship, increase production. For example, a man with a hoe can do more work than a man without one. The increase in production that has taken place in the United States through the use of capital goods is well known. Without expending any more energy, labor nearly always can produce more goods when capital is available.

This truth, however, although demonstrating the value of capital instruments, does not explain why capital in production earns not only enough to perpetuate itself but also something more that can be paid to the owner for its use. This question was posed years ago as follows: Suppose a person owns a new plane which he lends to a carpenter. Suppose the carpenter uses the plane and with it becomes more productive. Assume further that the carpenter wears out the plane. When asked to return the plane to its owner, he buys a new plane and returns just what the owner gave him, namely, a new plane. Since the owner gets a new plane as good as the one he lent the carpenter, how can it be

argued that the carpenter, besides returning the plane, should pay something for its use?

The answer lies in the fact that, since the plane made the carpenter more productive, owners of capital instruments will insist in sharing in that productiveness; otherwise they will not allow others to profit while using their capital. The owner of a capital good, for example, is deprived of the use of that good while it is being used by someone else. The user becomes more productive; hence he can afford to share with the owner the extra fruits of his labor.

It must be reiterated that the manner in which capital is used in combination with other factors determines the rate of interest that the producer can afford to pay. Under perfect competition and perfect mobility of the factors of production, each factor will be employed in its most productive use, and increasing amounts of each factor will be used in any productive process until the addition to total income no longer exceeds addition to total cost. At this point the payment to each factor will be just equal to the productivity of the last or marginal unit.

A businessman will continue to use capital in combination with the other factors until the increase in his total returns from the last unit that he employs just equals what he must pay the owner of capital for using it. We know from our study of the law of diminishing returns that the addition to total returns eventually will grow smaller and smaller as equal successive units of any factor are applied to a fixed combination of the others. Hence the product at the margin of successive units of capital will decrease, and a curve showing the trend of the value of the marginal product would be a curve falling to the right. Such a curve is the demand curve for capital. As the student will see, our explanation of the demand for capital is an application of the marginal productivity principle developed in Chap. 20, applied to the problem of rent in Chap. 21, and to wages in Chap. 22.

As was the case in employing labor, the businessman naturally wants to increase his use of capital as long as each additional unit adds increasing amounts to total returns, and his demand for capital will continue until the amount that he can afford to pay for capital will just equal the smallest amount that lenders are willing to charge for lending him capital.

In summary, capital will be demanded because, when used with other factors of production, it increases total productivity. The businessman must estimate the contribution of each factor, and he will use additional

capital as long as he is convinced that the productivity of each additional unit is sufficient to give him a return at least equal to the amount that he is forced to pay for it.

The student should note that all the units of capital employed will receive a return equal to the productivity of the last unit added. However, marginal productivity cannot be said to determine the interest rate, for this is dependent on both the supply of capital and the demand for the products of capital.

Supply of capital. We have just noted that those who produce will employ capital and are willing to pay for its use because it increases production. This explains why interest can be paid. Now we ask, Why must interest be paid? As we shall see, businessmen must pay interest to obtain the use of capital because effort is required to create it and it is scarce. In our discussion of the supply of capital, we must explain, therefore, how capital is created. As we mentioned earlier in this chapter, the supply of capital is derived either from the savings of individuals and business institutions or from credit created by commercial banks. Let us first consider that portion of the supply of capital that arises from the savings process.

Three things must take place before a return can arise from capital originated through saving. (1) Someone must give up present consumption in order to save money to pay workers. (2) Workers must toil and sweat to produce the capital instruments. (3) Once the instruments have been created, they can be used only during the passage of time, and hence those who have furnished the funds must wait before returns can be received.

To illustrate the creation of capital in this manner, suppose that an individual desires to invest $100 in a certain business. He has no funds and so he must begin by saving some of his wages. He decides to save $10 a month. In the first place he must forego certain pleasures in the present if he wishes to save money. The first process, then, is working for a wage and then saving part of that wage by abstaining from some present consumption. After 10 months, the individual has saved the $100. If he wishes to use capital goods himself, he must either pay workers to produce them or purchase them in the market. Otherwise to invest his $100, he must allow someone else to borrow it. In either case, before he can secure any return for the services of this capital, he must wait another period, say 6 months or 1 year. We see, therefore, that two sacrifices have

been made: (1) the sacrifice of abstaining from consumption during the period in which the money was being saved and (2) the sacrifice of abstaining from consumption when the capital instruments obtained with the money were being used for further production. It is clear that in both cases a waiting period was involved.

The rate of interest may be regarded as compensation for both these periods of waiting, though whether the rate of interest was really what induced the saver to give up present consumption is open to some doubt. It is safe to say, however, that some saving does occur, principally because people hope to get interest. We shall assume that, in the illustration given, the individual was induced to save by the fact that he would receive a certain rate of interest. The rate of return, therefore, must have been sufficient to compensate him for both types of sacrifice.

Foresight. Because of foresight, many individuals will save even without thought of securing a return. As a result, funds are often accumulated by individuals who are more interested in increasing and preserving their capital than in enhancing future income by investment. Under these conditions safety of principal and perhaps liquidity are primary requirements. Businessmen who could assure safety and liquidity would be able to borrow such funds even if the interest rate were very low.

Some funds are also accumulated and reinvested by corporations whose directors have decided that the interests of stockholders will be best served that way. Then, too, individuals may inherit great amounts of capital; sometimes so much that it is hard to spend the annual income. Some capital therefore is accumulated without effort or sacrifice on the part of those who save it.

On the other hand, many individuals never save. Some persons have a very low degree of foresight. The present is so much more significant to them than the future that special inducements must be given before they will save any of their incomes. Again, there are other individuals for whom the pressure of present needs is overwhelming because they have very small incomes. Such persons are never able to save, or if they save at all, they do so only at great sacrifice. Unless some very attractive inducement is held out to them, they cannot be expected to save. These are special cases. However, it is often maintained that there is a general tendency on the part of human beings to rate the present above the future; and to this tendency the name *time preference* is applied.

If we assume that interest must be paid in order to induce people to save, we have a simple answer to the question, Why must interest be paid? The answer is that interest must be paid to get individuals to accumulate capital. Some capital funds may be saved, of course, without much sacrifice. However, if such funds are insufficient to meet the needs of those who are willing, if necessary, to pay for the use of capital, inducements must be held out to offset the sacrifices of those who have so little foresight that they will not save unless given enough compensation to overcome their inclination to spend in the present.

The question often arises whether it is any longer necessary to depend for capital upon those who will save only in response to the inducement of high returns. Low rates of interest now prevailing offer much evidence that, in this country at least, the sources of capital supply are now so abundant that inducements of this nature are no longer necessary. Nevertheless, if in the future demands should arise for larger amounts of capital, higher rates of interest would prevail, and it is possible that these higher rates would cause a substantial increase in the amount of saving.

When individuals accumulate capital funds, it is clear that some sacrifice is usually involved. Is this equally true when funds are created by the banks through the extension of deposit credits to borrowers? Such extension of credit may involve no sacrifice for the bank. The borrower, however, must still make the sacrifices involved in saving. For him these sacrifices are merely deferred. In order to pay back the loan, the borrower must accumulate funds. The businessman is always faced with two alternatives: he may either borrow capital or accumulate it through his own saving. If he borrows, he must repay the loan through saving in the future. To accumulate capital, he must save in advance of the time when he needs it.

In answer to our question, then, Why must capital be paid? we discover that saving and investing entail real sacrifice for many individuals. Capital is therefore scarce except at high rates of interest. We discover further that an increase in demand will raise the rate of interest, and that a higher rate may bring a substantial increase in saving. A curve showing the relationship of the supply of capital and the rate of return would therefore be a rising curve. In other words, larger and larger amounts of capital would be called forth as higher rates of interest were offered.

Determination of the interest rate illustrated. We have explained the demand and supply curves for funds in a certain country in a certain

year. We have said that under conditions of competition the rate of interest will be that rate which equalizes supply and demand. We can illustrate the determination of the rate of interest by means of the chart [1] in Fig. 1.

The demand curve in Fig. 1 shows that those who desire capital will borrow more at low rates than at high rates. As they employ more capital,

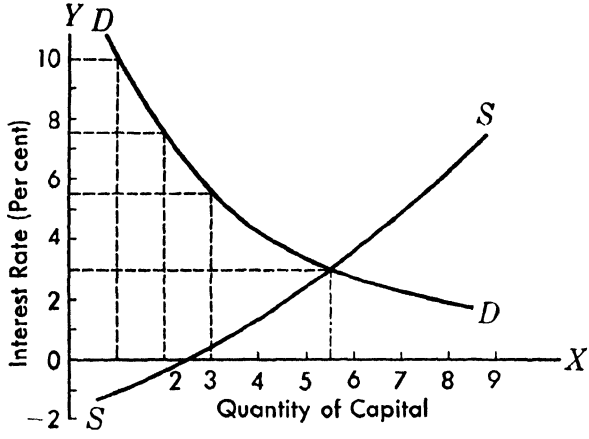


Fig. 1.—The supply and demand for new capital in a certain country in a certain year.

its marginal productivity drops, but the product may still be sufficient to cover a low interest rate. If only 1,000,000 units of capital are used, it will produce in combination with other factors a sufficient income so that those who use the capital can afford to pay 10 per cent. This is because capital under these conditions will be used only in very productive ways. If, however, 2,000,000 units are employed, a return of only 7½ per cent will be attributed to the services of capital. This conclusion follows because capital must be put to other than its highest uses, and the law of diminishing returns operates. When additional units of capital are

[1] The student should note that this chart refers to the supply of *new* capital and the demand for *new* capital. The total stock of capital on loan in a country at any given time is an accumulation from the past and, since it is already produced, would be offered for use at almost any price. Since this total amount would be very little affected by the savings of any one year, a curve representing the *total* supply of capital would be an almost vertical line. The demand curve for new capital in any one year should be highly elastic, since the amount of capital added in any one year could not greatly change the marginal productivity of capital.

employed, each unit being employed in the highest use still open, the chart shows that the marginal productivity of capital decreases.

On the supply side, a different situation is represented. The supply curve shows that some capital funds would be accumulated and lent even though the people who accumulated such funds had to pay something for their protection. Here the degree of foresight is so great that even part of the principal will be sacrificed to protect the rest. However, this chart is based on the assumption that, under such conditions, the amount of capital saved would be small. The chart also shows that some individuals would lend capital without any return so long as their principal remained intact. Here also safety of principal is the prime consideration. Beyond this point, rewards must be offered if individuals are to be induced to save and invest. Moreover, the importance that people attribute to future consumption will determine how much they will save at each rate.

Under perfect competition and perfect mobility of the factors of production, those who desire capital will determine how much they can afford to pay for such capital, and those who save capital will determine how much they are able and willing to accumulate at various rates of interest. Those who demand capital will decide what the demand schedule is to be, that is, the quantity that will be taken at each possible rate. So also those who supply capital will determine the supply schedule. The point at which the demand and supply schedules meet will establish the price of capital, or the interest rate.

Thus we see that the demand for capital is largely determined by its marginal productivity when different amounts are employed, and if we assume that interest is what induces people to save, the supply of capital is largely governed by the interest rates that must be offered to marginal savers to secure various amounts of capital. Under normal conditions, if capital is relatively scarce, its productivity will be high and businessmen will be willing to pay a high rate for it. If capital is relatively abundant, its significance in the productive process will not be so great, and the rate that businessmen will pay will be less. The interest rate will always tend to equal the marginal productivity of capital. However, both the interest rate and the marginal productivity of capital depend on the relation between supply and demand; and both will be established at the point where supply and demand are in equilibrium.

In the above discussion we have taken the traditional point of view

with respect to the supply of capital. This view is that the receipt of interest is the principal inducement for saving, and that therefore the amount saved depends on the interest rate. However, the assumption that the interest rate is the principal factor that influences saving cannot be proved and, as already suggested, is at least open to question. Some economists believe that the interest rate has only a minor effect on the volume of savings, and that the principal inducements to saving are of a different nature. If they are right, then the supply curve for new capital in Fig. 1 should be represented as much more inelastic than it is, to indicate that the amount of capital funds offered for investment is not much affected by the rate of interest.

Lord Keynes, the famous English economist, emphasized another point, namely, that saving money and investing it in capital goods are two distinct processes. He believed that a great deal of saving might take place independent of the interest rate; but if the interest rate fell too low, the funds saved would be hoarded rather than invested. In such a case the function of the interest rate would be not to induce saving but to induce savers to give up liquid holdings of money in exchange for somewhat less liquid investments.

Interest rate and the quantity of money. Since capital is often regarded as identical with money, the statement is frequently made that, if the amount of money in circulation were increased, interest rates would be lowered. This theory has been advanced by those who have advocated cheap money policies. It was explained in an earlier chapter that in the long run the only effect of an increase in the quantity of money in the economic system is to increase prices. Although no exact relationship exists, it is known that increases in the amount of money are accompanied in the long run by increased price levels. If this is the case, then an increase in the amount of money cannot in itself cause a decrease in the interest rate in the long run.

Of course, an increase in the amount of money may temporarily cause a decline in interest rates, because the supply of funds may increase faster than the demand for them. However, with every increase in prices, more money will be required to produce or purchase a definite quantity of goods. Consequently, producers must secure a greater amount of money to purchase capital instruments. Consumers indeed may have more dollars in their possession and therefore may be willing to lend at a lower rate an amount of dollars equal to that which they were willing

to lend before the quantity of money increased. However, after allowing time for the price level to become adjusted, the increase in the demand for funds, resulting from higher costs of production, will tend to offset the increase in the supply. We may conclude therefore that in the long run the pure interest rate will not be changed. Long-run changes in the pure interest rate may result from changes in the desire to save, or changes in the productivity and the demand for capital, but not from changes in the quantity of money.

When competition is not perfect. So far in our discussion we have assumed that perfect competition prevailed in the market for capital. Actually, this is not true, but the absence of perfect competition does not nullify our analysis. We can still say that the interest rate is determined by supply and demand, and most of the differences that we observe in particular rates are due to differences in risks and costs. It is true that in any one area of a large country, such as the United States, there may be relatively few lenders compared with the number of borrowers. Hence lenders may wield some degree of monopolistic power and temporarily may obtain a higher rate of interest than would prevail under perfect competition. Such differences, when loans of equal quality are involved, probably will not persist for long because of the growing number of lending organizations, such as insurance companies, which operate on a nation-wide basis. Lack of knowledge among borrowers in regard to interest rates may also enable lenders to obtain more than would otherwise be the case. This also is temporary. In short, the absence of perfect competition may account for greater differences in contractual rates than would occur in the highly competitive situation that we have assumed in our analysis. Such differences, however, do not seriously affect the validity of our explanation of interest rates in terms of supply and demand.

Summary

Interest, or the share of income that accrues to capital, has been explained by supply and demand analysis. Interest is a payment made each year to the owners of capital, and we customarily express it as a rate of return. This rate of return is a percentage which can be computed by dividing the annual payment for the services of capital by the value of the capital.

In using demand and supply analysis, we said that the rate of interest in a certain country in a certain year is that rate which will equalize supply and

demand. At this rate, known as the equilibrium rate, the amount of capital wanted by borrowers will be equal to the amount that lenders are willing to supply.

Demand and supply analysis explained only the pure rate of interest, which was defined as the rate of return on capital after all expenses and compensation for risk had been deducted. We used the term contractual interest to designate the lump sum annual payment that the lender receives from the borrower. This payment includes compensation for all risks involved in lending and for all expenses and costs incurred. Differences in contractual rates can be explained by differences in costs, expenses, and risks.

The demand for interest, we learned, came for the most part from businessmen who needed capital for productive purposes. Businessmen have an incentive to borrow, we saw, as long as the increase in total returns from use of capital is greater than the increase in cost from acquiring and using capital. Since the law of diminishing returns comes into play as more and more units of capital are used, the businessman will pay less and less for additional units of capital. Thus we know that the demand curve for capital slopes downward to the right. The supply of capital comes in large part from individuals who can be induced to defer consumption. This process of waiting may involve saving or merely deferring consumption of funds already accumulated. Since the inducement to defer consumption probably is greater the higher the interest rate, we can assume that the supply curve for capital rises to the right. The equilibrium rate of interest can be illustrated by intersection of the demand and supply curves.

The student was warned to remember that interest cannot be determined independently of the return paid to the other factors. The businessman who borrows capital also employs other factors. He compensates all factors by selling a product. The most that can be paid to capital, as well as any other factor, is the contribution of that factor to production.

In the next chapter, we shall consider profits, or the share of product that goes to the entrepreneur.

Study Questions

1. Define capital, interest, wealth, liquid capital, free capital.
2. Why is interest expressed as a rate and not an amount?
3. As the term interest is commonly used, what other factors besides a return on capital are generally included?
4. How is the rate of interest determined for someone who wishes to use the borrowed money for consumption purposes?

5. Distinguish between the contractual and the pure interest rate. Which rate is of more concern to the authors in the text?
6. Just how is the interest rate determined?
7. What two questions must be asked when the rate of interest is to be found?
8. What is meant by the statement "Money is barren"? Is this correct? Explain.
9. Justify the payment of interest.
10. How would you find the demand curve for capital?
11. What three things must take place before a return on capital can arise?
12. How does foresight affect the interest problem?
13. Since banks create funds by the extension of deposit credit to borrowers, are not such funds created without any sacrifice? Explain carefully.
14. How do the forces of demand and supply finally determine the point of equilibrium that represents the interest rate?
15. What relationship, if any, is there between the interest rate and the quantity of money? Explain carefully.
16. Does the fact that competition is imperfect seriously affect the determination of interest rates? Why or why not?

Exercises

1. An investor purchases a plot of land for $300. He receives an income of $15 per year from the land. Is this return rent or interest? Justify your answer. Suppose that the average rate of return in this community becomes 4 per cent and suppose the income received by this man increases to $24 per year from the land. Would the value of the land be the same as above? Why or why not?

2. A man decided to save a sufficient amount of capital to allow him to receive an income of $200 per month.

a. If the rate of interest is 5 per cent, how much must he save?

b. Suppose that the rate drops to 3 per cent. Will this affect the amount he must save?

c. Under these conditions does not the man save more at low rates than at high rates? Does this not show that the principles explained in this chapter are in error? Explain carefully.

3. Interest rates are very low at present. Would you agree with the following conclusions?

a. Capital is unproductive; therefore little can be paid for the use of this factor.

b. Individuals save for other reasons than to secure a rate of return; therefore interest rates can be low.

c. Banks can create credit; therefore it is not necessary for individuals to make sacrifices.

d. There is so much money in circulation that no one has to borrow.

4. The statement is made that, even if banks can extend credit, sacrifices must still be made, for the borrower must make them to pay back the funds. Thus these sacrifices are merely deferred.

If this condition prevails, should we not pay the borrower for making the sacrifices rather than pay the bank for extending credit? Explain.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 19.

Bohm-Bawerk, E. von: "Capital and Interest," Brentano's, New York, 1922.

Boulding, K. E.: "Economic Analysis," Harper & Brothers, New York, 1941, Vol. I, Chap. 14.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 21.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 20.

Fisher, I.: "The Rate of Interest," The Macmillan Company, New York, 1930.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 17.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 26.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 28.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 29.

MOULTON, H. G.: "The Formation of Capital," Brookings Institution, Washington, D. C., 1935.

SPAHR, W. E., and OTHERS: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chap. 29.

SLICHTER, S. H.: "Present Savings and Postwar Markets," McGraw-Hill Book Company, Inc., New York, 1943.

TAUSSIG, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. II, Chaps. 38-40.

24. Profits

So far in our study of distribution we have described the payments made to land, labor, and capital for their contributions to production. We have one more share to discuss: the payment to the entrepreneur who initiates and directs production. We use the term *profits* to designate his share.

Identification of the entrepreneur and segregation of the payment for his services from other shares may be difficult in some of the complicated forms of modern business. This is especially true of the corporate form of business organization. A good example of an entrepreneur, however, is the businessman whom we often have mentioned in this text. Because there is some misunderstanding of the entrepreneur and his contribution to production, it is well to describe his functions thoroughly before discussing profits.

Role of the entrepreneur in a free society. In a society built upon the principle of free enterprise and personal liberty, the entrepreneur is a very important person. If we become convinced that socialism, communism, or fascism is a better form of society than the free-enterprise system, we may decide that we can do without the entrepreneur or that the state should regulate his activities down to minute details. But in a free-enterprise system, the entrepreneur is even more than an important functionary; he is a symbol of a free society. He is the dynamo of the business process; he initiates and conducts the enterprises that provide employment for the other factors.

Let us begin our study of him by considering the case of a grocery store, which a Mr. John Jones starts in a small Middle Western town. First of all Mr. Jones had to have the idea for the business. At some time he decided that a grocery business would be successful in this particular town. Next he arranged for the financing. He had some money of his own to invest and he borrowed more. He assumed a personal obligation to pay back the funds borrowed.

With the capital at hand, he selected a site and signed a lease with

the landlord. He agreed to pay a monthly rental for a period of years. Then he invested some of his capital in remodeling the store that he had rented, in fixtures, and in a stock of groceries. Next he hired several clerks and opened his store. Under his careful supervision the enterprise prospered because it performed a useful service in the community.

We now should ask ourselves two questions, Why did Mr. Jones go into business for himself instead of going to work for somebody else and being assured of a regular pay check? What services has Mr. Jones performed for society?

Several answers might be given to the first question. Mr. Jones might be an independent, venturesome fellow who would rather work for himself than for someone else. If Mr. Jones is so constituted, we are all for him. A progressive society must have venturesome individuals with ideas and initiative.

There probably was, in addition, a pecuniary motive. Mr. Jones hoped he could do better for himself financially by operating his store than by working for someone else. He expected the store to make a profit, and he hoped the profit would be larger than the salary he would have been paid as a hired store manager. The hope of large profits, then, was one of the reasons for starting the business.

This brings us to the second question, What services has Mr. Jones performed for society? Do these services entitle him to a greater reward than he could have enjoyed working for someone else?

Let us review what services Mr. Jones performed when he opened his store. We might list them as follows:

1. He conceived the idea for the enterprise.
2. He initiated the enterprise; that is, he opened the store.
3. He performed the service of management.
4. He assumed certain risks and responsibilities: (*a*) the responsibility of paying back the money he borrowed; in addition, he risked some of his own capital. (*b*) He signed a lease which obligated him to make monthly payments to the landlord. (*c*) He probably contracted debts for remodeling and stocking the store. (*d*) He assumed the responsibility of paying wages to his employees; he agreed to "meet the payroll." (*e*) He assumed certain responsibilities to his community, such as paying various business taxes, collecting pay-roll and social security taxes from his employees, and sharing the burden of making his community a better place in which to live.

If Mr. Jones is to remain in business, he must be rewarded for the services he performs. There must be some rate of compensation that would just induce him to operate his grocery store, just as there is a rate of wages that will just induce one of his clerks to come to work every day. Mr. Jones receives his compensation in the form of profits, of course, so we can say that he must earn sufficient profits over a span of years to make him willing to stay in business, or he will go to work for someone else or enter another occupation. We can use the term *wages of management* to indicate that rate of profit that will keep Mr. Jones in business. The term *normal profits* also could be used to express the same idea.

If Mr. Jones does not receive normal profits, or wages of management, whichever term we care to use, he will quit business. From year to year, his profits may vary because the tempo of business is faster in some years than in others. He knows that he can expect that. However, he will elect to close his store if profits seem to be on a long slide from which there is no hope of recovery.

What is the effect on others if Mr. Jones does not earn normal profits? If he decides to go out of business, his workers lose their jobs, the landlord must find a new tenant, and the capitalist must look elsewhere for a borrower. There is another effect. The services provided by Mr. Jones's grocery store no longer are available. Since this happened because Mr. Jones did not earn normal profits, we can say that normal profits are a cost to the community of obtaining the services of Mr. Jones and his enterprise. Obviously normal profits are no different from any other costs that must be met in the long run if a good is supplied.

When Mr. Jones decided to go into business, he beyond doubt had something more in mind than merely wages of management. He hoped to make profits well in excess of this figure, and it probably was the lure of this something extra that decided him to assume the risks of business for himself instead of accepting a pleasant, easy job as the paid manager of a grocery store owned by someone else. Profits in excess of normal profits or wages of management can be called *pure profits.* In boom years, *pure profits* probably will be earned by the grocery store, and it is possible that Mr. Jones will be so skillful in running his business that pure profits will be generated even in ordinary years.

However, we can say that, in the long run under perfect competition, there will be no pure profits and only normal profits will be earned. Why? Because the existence of pure profits will attract competitors into the

grocery business, as into any industry where profits are above normal. The presence of more firms will force prices and profits back to normal levels. Contrariwise, if profits within an industry fall below normal for long, the withdrawal of firms will force earnings back up.

It is quite possible, too, that Mr. Jones will in the future regard a higher rate of earnings as normal if he is to remain in business. When he once demonstrates his ability to run a grocery store perhaps more efficiently than other proprietors in his locality, he will insist on being rewarded accordingly. If his community is not able to support a store that will earn this higher level of profits, Mr. Jones may decide to move somewhere else. Hence normal profits might now include dollars that formerly would be classified as pure profits.

Other payments that may be included in profits. Risk is involved in any business enterprise, of course; but the greater the risk, the larger the possible profits of the entrepreneur must be. Thus we would expect larger profits to be earned by enterprises where the danger of disaster is always at hand. One of the important risks of the entrepreneur results from production in anticipation of demand. Those who correctly anticipate future demand are in a position to secure commensurate profits. Those who anticipate demand incorrectly will suffer losses. The degree to which production must be carried on in anticipation of demand will influence what entrepreneurs regard as an adequate rate of return in a particular industry.

Many students of business have observed, however, that some business concerns regularly have larger profits than others. The difference cannot always be accounted for by the risk involved, because sometimes substantial differences occur within the same industry. How can this be explained?

If profits exceed the amount necessary to induce the entrepreneur to perform his services and assume the necessary risks and responsibilities, the excess may be due to imperfection of competition. This may be openly sanctioned by society; for example, the enterprise may control important patents. Sometimes the excess profits are due solely to the superior skill of the entrepreneur. A number of other reasons might be cited for differences in earnings between industries or between firms within the same industry. Sometimes the real reason for such differences may be difficult to determine.

Relationship between risk and profits. The relation between risk and the receipt of profits has often been misunderstood. It is sometimes said that employers secure profits because they take risks. This causal relationship, however, does not usually hold.

The relationship between risks and profits is as follows: In the hope of securing profits, employers take risks. Furthermore, because of the risk involved the number of employers will be limited. When the number of units of any factor is limited, marginal productivity ordinarily is higher than when the number of units is great. Consequently, to the extent that danger of failure because of the risks involved limits the number who enter a certain field, successful businessmen will receive higher rewards than they could get in industries where risks are less.

Those who take risks, then, are performing a great social function. Employers must anticipate a need; they must discount hazards; they must have vision and foresight. The great pioneer captains of industry often made great profits. They were noted for their willingness to bear risks because they believed that their ideas for new railroads and new factories would fill a need and so yield profits. Thus, those willing to take risks often secure high profits, although it does not follow that businessmen who take risks are guaranteed profits.

Two qualifications of this analysis must now be noted. First, all risks are not borne by employers. So far as future risks can be foreseen, they may be shifted. Thus the chances of fire, flood, and the like may be anticipated and the risk of loss shifted by purchase of insurance. Some businessmen shift as many risks as possible to others in the marketing process or to the consumer. Even so, after all insurable and other anticipated risks are shifted, some remain, and these must be borne by the businessman.

A second qualification is that all risk bearing is not socially desirable, and hence profits that arise from certain kinds of risk are not considered socially desirable. The gambler, for example, takes a risk at dice, at the race track, and the like. Yet his gain is simply loss for someone else. Society has not benefited.

Distinction between the use of the term "profit" by the accountant and by the economist. The accountant and the economist use the word profit in a somewhat different sense. The accountant says that the net income that accrues to owners of a business is profit. Hence, the accountant regards profits as a residual after all payments to land, labor, and capital, and after payment of taxes. The economist, on the other

hand, regards profit as the share of product that goes to the entrepreneur for his special services. Since the entrepreneur performs useful services, he must be compensated if he is to continue to function. Hence the economist regards normal profits or wages of management as a cost, just as payments to labor, land, and capital are costs. Society must compensate the entrepreneur if it is to enjoy his services. In summary, the accountant looks upon profits as a residual after all costs are met; the economist regards normal profits as a cost. However, anything that accrues to the entrepreneur over and above normal profits the economist calls pure profits, and these he does not regard as a cost.

Profits earned in American industry. A study of the record made by business in general would indicate that it is not so easy as the outsider may think to make a profit in an accounting sense. In fact, many misconceptions exist about profit margins, which from the accountant's point of view are what remain after all expenses and costs are met. For the most part, profit margins generally are believed to be much higher than they are. In boom years, profits usually do become substantial, but heavy losses are incurred by many business concerns during depressions. A fair analysis of profits would call for an inspection of the average level of profits or what profits were for the full span of the business cycle rather than for the individual years. It has been customary, however, to compute profits on an annual basis, and this also is necessary for purposes of taxation. Hence wide swings occur in the rates of profits earned from year to year.

The over-all level of profits earned by business in any one year also may not tell the whole story. As an example, corporation profits in 1946 were $12.5 billion after taxes. But in this highly profitable year, the most lush in the history of American business up to that time, some companies were still in the red or would have been except for tax carrybacks. Even in 1947, when 1946 corporate profits were exceeded by $4 or $5 billions and when postwar reconversion problems had been solved, some companies reported losses.

Indicative of how accounting profits fluctuate, the $12.5 billion level of profits in 1946 compares with average annual profits for the years 1936 to 1939 of a little less than $4 billion. A still more complete picture of how profits fluctuate can be had from Table 30, which gives a percentage distribution of manufacturers' costs in the period between the world wars. This table deals with manufacturing alone. It is interesting to note the

very much greater fluctuations in profits than in cost of materials, wage and salary payments, and in other costs.

Table 30. *Distribution of Manufacturers' Costs in the Value of Production,* 1919-1939 *

	1939	1937	1935	1933	1931	1929	1927	1925	1923	1921	1919
Production value	100	100	100	100	100	100	100	100	100	100	100
Cost of materials..	36.9	39.8	41.2	35.5	34.0	34.4	36.5	36.1	35.7	40.1	39.2
Wages—salaries ..	30.8	31.8	31.5	33.5	36.2	32.6	33.4	33.1	35.0	35.3	32.8
Other costs.....	28.2	24.8	25.3	33.8	36.1	28.2	27.6	26.6	24.9	29.0	23.3
Profits..	4.1	3.6	2.0	−2.8	−6.3	4.8	2.5	4.2	4.4	−4.3	4.7

* SOURCE: The Conference Board.

The actual dollar level of profits and losses resulting from manufacturing activity alone in these same years is given in Table 31.

Table 31 *

Year	*Amount (millions)*	*Year*	*Amount (millions)*
1919	$ 1,917	1931	$−1,837
1921	−1,323	1933	−587
1923	1,777	1935	646
1925	1,742	1937	1,498
1927	1,067	1939	1,595
1929	2,319		

* SOURCE: The Conference Board.

The large number of business failures even in normal years gives a grim picture of the risks involved in business and proves conclusively that accounting profits do not accrue automatically merely because an enterprise has its doors open. According to Dun and Bradstreet data, there was a monthly average of 1,231 failures in the United States in

1939; 1,135 in 1940; and 987 in 1941. During the war years and early postwar months, economic activity reached such feverish heights that failures dwindled to record low levels. Only 42 failures were reported in December, 1945. But experienced businessmen knew that such favorable conditions would not last.

The large year-to-year fluctuations in profits in an accounting sense mean that the income of the entrepreneur is not so stable as the income that accrues to the other factors. In one year he may have large profits and in the next year large losses. That occurred between 1920 and 1921 for many businesses, and for a substantial number between 1937 and 1938. The large fluctuations make the concept of normal profits in an economic sense more nebulous than normal wage rates.

The concept of normal profits, however, is a useful one. In some years the businessman will make less than normal profits; in other years he will receive more. The year-to-year profits fluctuate around the normal profits. It is the normal profits that he considers when deciding whether he should start an enterprise or continue one which he has already established.

Criticism of profits. Critics of the profit system often contend that the government should assume more responsibility for smoothing the booms and depressions that increase the risks and uncertainties of operating a business. They argue that government planning and regulation of competition could decrease the responsibilities and burdens of businessmen. By careful planning in regard to the number and kinds of businesses needed, and by control of the prices of raw materials and of finished products, they claim that the government could perform many of the services for which the entrepreneur now is paid.

Other critics of the profit system attack from another angle. They start from the premise that the wages of management cover all payments that need be or should be made to the entrepreneur if competition were perfect. They say that the entrepreneur, through monopoly power, exacts tribute from society which should never be collected or which represents shares of product belonging to other factors. Thus a charge of exploitation of society or of the other factors often is hurled.

The first criticism is concerned primarily with the kind of economic system we desire. If we want the government to take over the functions of the entrepreneur, we must accept state socialism or some form of government other than we now have. We also should admit our willingness to sacrifice personal and political liberty as well as economic free-

dom. If the state assumes the responsibilities now accepted by the entrepreneur, the state must have the right to decide which businesses may operate and which may not. The right to hire and fire the paid managers of any enterprise would have to rest with the state. This privilege of necessity would extend even deeper. The government would have to keep for itself the right to force labor to work at whatever kind of work was regarded as vital, regardless of the wishes of the individual. Powerful unions could not, of course, be tolerated because they might impede what the government wanted. Economic liberty and personal freedom appear to be linked together.

The second criticism is more difficult to answer. However, regulation of monopolies by Federal, state, and local governments has done much to prevent exploitation of the consumer. As we shall see in Chap. 27, consumers can aid in preventing exploitation by insisting upon complete information on products offered to them. Society can promote competition in our economy by maintaining a favorable atmosphere for free enterprise. This would include a well-conceived tax system, labor laws equitable to both employer and employee, and freedom for businessmen from excessive regulation.

Entrepreneur in a corporation. As mentioned earlier in the chapter, the identification of the entrepreneur may become difficult in modern forms of business organization. Who is the entrepreneur in a large corporation? The corporation has a president, who receives a salary, and a board of directors, who devote part time to the enterprise in return for fees and the prestige involved. The board of directors selects the president and the other officers, or delegates selection of the other personnel to the president.

Under such circumstances, the real entrepreneurs are the stockholders. The housewife who owns a few shares of General Motors stock may be startled to learn that she carries such a responsibility. The stockholders of a corporation elect the directors and delegate to them the responsibilities of entrepreneurship.

The directors may in turn hand over the functions of entrepreneurship to the president and other officers, who really are hired managers and receive wages, although many corporation officers also are stockholders.

The ultimate responsibility rests with the stockholders, however, whether they accept it or not. They have the right to attend stockholders' meetings and to question the policies of management. It may be difficult

to exercise this privilege. The better corporations try to enlist the interest and patronage of stockholders, but often they have difficulty in getting more than a corporal's guard at meetings. Corporations are required to send notices of stockholders' meetings and proxies for those who cannot come to sign so their stock can be voted. Many proxies are not returned, and the stockholders do not attend. It seems to be a question of apathy. However, there are numerous instances where stockholders have risen up and ousted a management that did not appear to be performing its duties with skill.

The stockholder, as an entrepreneur, seems to be very different from Mr. Jones, the entrepreneur who founded the grocery store in a small Middle Western town. The typical stockholder probably did not initiate the enterprise in which he has stock; instead he purchased shares in a going concern. He is very much interested in the profits that the corporation earns but he does not expect to receive all these profits. He can hope to have only his proportionate share of that part of profits which the directors decide to pay out as dividends. The stockholder may have purchased his shares in hope of capital gains through price appreciation rather than because of the dividends that may be declared.

Despite all these differences between Mr. Jones and the stockholder in a large corporation, the stockholder nevertheless is an entrepreneur. Whether he likes it or even believes it, he assumes the risks and responsibilities for the corporation and with other stockholders he exercises the ultimate control. He profits through larger dividends and capital appreciation if it is successful. He loses the price he paid for his share if it fails. Any corporation will be more successful if the stockholders are conscientious in the exercise of their entrepreneurial functions by checking up regularly on the performance of their managers

Officers of corporations often are major stockholders of small and medium-sized corporations. Ownership of big corporations usually is diffused among many thousands of persons. When officers are newly employed, the stockholders frequently encourage stock ownership by making the acquisition of shares easy and advantageous. For example, a new president may be given an option to purchase a bloc of stock at a certain price. The option is good for a number of years. If the president does well in managing the company, the price of the stock may increase far above the option price. Thus the president has a pecuniary incentive, in addition to that provided by his salary. When he owns stock

or can acquire it under the option, he becomes an entrepreneur as well as a hired manager.

Because stockholders supply capital to corporations through purchase of shares, they do more than perform some of the entrepreneurial functions; they also are capitalists in a sense. There is an important difference, however, between the interest payment received by the lender of capital and the dividends or capital gains that accrue to the stockholder. The lender of capital has no right to the assets of the corporation unless it defaults in its payments to him; the stockholder owns a proportionate share of the assets. In the case of failure, however, the lender has a preferred claim. The lender is entitled to a specified rate of return while the stockholder gets only what the success of the corporation justifies. The position of the lender is much more secure than that of the stockholder. In short, the stockholder is an owner who assumes the risks of ownership. The lending capitalist merely has made a loan and is entitled to repayment thereof when due and to interest in the meantime.

Summary

We have described in detail the payment to the entrepreneur of what we call profits. Profits are the earnings of the entrepreneur. They are the payments made to him for initiating, directing, and bearing the responsibility for a business enterprise. We used the terms *wages of management* and *normal profits* to indicate that rate of earnings which would just compensate the entrepreneur for his services. Any profits in excess of this amount were called *pure profits*. We saw that there will be no pure profits in the long run under perfect competition. However, *wages of management* must be covered in the long run if an enterprise is to function, just as wages of labor, capital, and land must be met.

We have stressed the point that the entrepreneur is not only a very important person in an economic system based on free enterprise. He is also the symbol of free enterprise. Those who want the government to take over all his functions are in reality asking for some other kind of economic system.

Study Questions

1. What is the role of the entrepreneur in a free society?
2. What do we mean by the term profits?

3. What functions did Mr. Jones perform for which he should be rewarded by receiving profits?

4. What is the meaning of pure profits? normal profits? wages of management?

5. Explain the relationship between risk bearing and profits.

6. What kinds of profits are socially undesirable? Why?

7. Distinguish between the use of the term profit by the accountant and by the economist. In what sense is each correct?

8. Has industry over the years earned high profits?

9. What criticisms have been advanced against profits? How are these criticisms answered?

10. To what extent and in what way is a stockholder in a corporation an entrepreneur?

Exercises

1. An editorial writer in the *Saturday Evening Post* in the May 3, 1947, issue made the following statement:

> Instead of regarding profit as a symbol of success in a free economy, as well as an index to the stability of wages and other forms of income, a lot of people have got it into their heads that profit is something which businessmen and corporations extract more or less feloniously from the economic process.

a. Does this statement agree with the point of view expressed by the authors in this chapter? Does it disagree?

b. Do you agree with this statement? Why or why not?

c. Is the writer talking about pure profits, normal profits, or residual profits?

d. The writer defines profits as follows: "A profit is what the man who runs a drugstore—or the corporation which operates 15,000 filling stations—has left over after paying wages, taxes, rent, interest, bills for raw materials, repairs, and other expenses." Does this definition of profits agree with that given by the authors? If not, in what respects does it differ?

e. Another definition given is: "Profit is the happy consequence of industry, good management, imagination or good luck, and it accrues to a surprisingly small proportion of those engaged in business ventures." Do you agree with this definition and its implications? Explain carefully.

2. It has been suggested that progressive taxes be placed on profits so that no individual or company could earn over a certain amount or percentage of invested capital.

a. What are the implications in such a suggestion as to the nature of profits?

b. What harm would result from such a procedure?

c. If you decided to limit profits, would you limit those profits which result from increased efficiency? Why or why not?

3. Explain carefully why our system of free enterprise, instead of being called a profit economy, should be called a profit-and-loss economy.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 21.

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 23.

Deibler, F. S.: "Principles of Economics," McGraw-Hill Book Company, Inc., New York, 1936, Chap. 22.

Friday, D.: "Profits, Wages and Prices," Harcourt, Brace and Company, Inc., New York, 1920.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chap. 18.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 29.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn and Company, Boston, 1937, Chap. 27.

Hardy, C. O.: "Risk and Risk Bearing," University of Chicago Press, Chicago, 1931.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chap. 32.

Knight, F. W.: "Risk, Uncertainty and Profit," Houghton Mifflin Company, Boston, 1921.

Marshall, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1925, Book VI, Chaps. 7 and 8.

Spahr, W. E., and Others: "Economic Principles and Problems," Rinehart & Company, New York, 1940, Vol. II, Chap. 31.

Taussig, F. W.: "Principles of Economics," The Macmillan Company, New York, 1939, Vol. II, Chaps. 48-50.

25. Taxation

The spectacular rise in the cost of operating governments has pushed taxation into the foreground as a major economic problem in almost every civilized country. This definitely is so in the United States, and it is a certainty that the economic well-being of every American will be influenced in the years ahead by the tax policies that are adopted by our various layers of government.

Many examples can be found in the United States and abroad of the higher cost of government. For instance, our Federal government spent only $725,000,000 in 1913, the last full year of peace before the outbreak of the First World War in Europe. In 1938, the last full year of peace before the Second World War, our Federal government spent $7,239,000,000 or roughly ten times as much. Expenditures by our states also increased nearly tenfold in the same period, and the cost of supporting our local governments was about four times as great in 1938 as in 1913.

Moreover, we see repeated in the period following the Second World War a lesson that history has demonstrated many times in the past: Wars step government expenditures onto a permanently higher plateau. There is no road back to the lower spending levels of the prewar period, as high as they may have seemed at the time. We saw this process after the First World War when Federal expenditures leveled off to $3,295,000,000 in 1923, as compared with the $725,000,000 spent in 1913. It now appears that Federal expenditures will fall in the 30 to 40 billion dollars range in most years following the Second World War, compared with the $7,239,000,000 spent in 1938 and the $8,707,000,000 spent in 1939. It would appear that the Second World War may have permanently raised our Federal expenditures almost as much percentagewise as the other war and very much more in terms of dollars.

Why has the cost of government risen? We already have indicated that war is in part responsible for the rising cost of government. Wars balloon national debts which must be serviced. Interest on our public

debt in the present postwar period has been running around 5 billion dollars a year. This alone is more than was spent to operate the entire Federal government in any peacetime year before the middle 1930's. Payments to veterans nearly always are boosted to a permanently higher level by a major war. The costs of maintaining a peacetime military establishment were spiraled skyward by wartime military inventions. Indeed, the cost of war does not end when the last cannon is fired.

There are reasons other than war and its aftermath that are to a major degree responsible for the rising cost of government. Citizens have been demanding more and more services from nearly all layers of government. When Adam Smith published his famous "Wealth of Nations" in 1776, he visualized only three main functions of government: protection against enemies from abroad, maintenance of law and order at home, and the provision of certain essential public services which private individuals could not undertake profitably. All this and very much more now are required of government. Recreation, old-age pensions, protection against the ravages of the business cycle, support of agricultural prices, destruction of rodents and other pests, preparation of bulletins on home canning, free public concerts, erosion control, and public informational services are only a few of the many new functions that now are demanded of government. The performance of these new services naturally has inflated governmental costs.

Even a quick survey of the complex system of government that has been set up in the United States would easily explain why government costs have become so high. A recent estimate placed at more than 165,000 the number of separate governing bodies with taxing powers in the United States. These included the Federal government, the states, counties, incorporated cities and villages, townships, school districts, and a variety of districts established to provide special services like parks or sanitation. The total cost of maintaining all kinds of government in the United States in 1939, the year war began in Europe, was nearly 17 billion dollars, equal to about 18 per cent of our gross national product.[1] In the years ahead, it appears that the total cost of all government may hover not far below the 50 billion dollar mark. This would be a little

[1] The gross national product is the value of all goods and services produced. It is greater than the national income. To derive the national income from the gross national product, it is necessary to subtract such items as certain business taxes, depreciation, and depletion.

under 22 per cent of the 230 billion dollar rate attained for gross national product in the early postwar period. Unfortunately, however, gross national product is likely to fall considerably below this 230 billion dollar level in many postwar years.

Sources of government revenue. The two principal sources of government revenue are taxes and loans. If a government is to remain solvent and does not wish to carry a large debt permanently, it must repay what it borrows. It can delay the repayment day by creating new loans to refund each old loan as it comes due. However, taxpayers must provide the means if loans are repaid permanently. In addition to receipts from loans and taxes, most governments receive at least some revenue from fees charged for services of direct benefit to specific persons. Some of these fees are for routine services, such as inspection of elevators, while others are for essential public-utility services usually provided by private companies. For example, many cities own their own power plants and charge for the services enough to cover costs and, in some cases, enough more to finance some of the other municipal services.

We can easily see the distinction between a tax and a fee. We can define a tax as a compulsory payment made by an individual to government without any direct *quid pro quo;* that is, the payment is not directly related to any service that the taxpayer receives from government. In the case of the fee, however, the taxpayer makes his payment and receives a direct service in return. The payment is directly related to the services; in the case of the tax, there is no such direct relationship. Fees are relatively unimportant in amount as compared with taxes, but fees are of greater importance to cities, towns, and other local governments than to the state or national government.

Characteristics of a well-designed tax system. Since such large sums must be collected each year from taxpayers, it has become of increasing importance that our tax system be as well designed as possible. What are the characteristics of a well-designed tax system? A satisfactory answer will give us bench marks for appraising our own tax system or those of other nations. Some characteristics of a good tax system will be accepted almost without argument by everyone; others will call forth much debate. We shall name three characteristics:

1. A well-designed tax system should provide adequate and dependable revenue and be capable of yielding increased revenue with the minimum disturbance to economic life in case of emergency. It also

should be capable of contraction with the minimum of controversy when the emergency passes. A tax system capable of providing adequate and dependable revenue probably should include a wide variety of taxes. Some taxes, such as those on incomes, vary considerably in yield from boom to depression. Others, such as taxes on articles of consumption used by the masses, are much more stable in yield.

The characteristic of adequacy and dependability would be approved in principle by most tax experts, but many economists would welcome wide fluctuations in tax yields between good and bad years. Heavy tax collections in boom years would cut down public spending power, these economists point out, and would thus reduce the danger of inflation. They further argue that low tax collections in years of low economic activity would be desirable even though the government might face heavy relief expenditures. Low tax collections, they contend, would leave more purchasing power in the hands of individuals when it was needed; and the reduction in taxes would not force the government to reduce expenditures since the government could meet its needs by borrowing. By spending borrowed funds, the government would actually add to the sum total of purchasing power.

2. A second characteristic of a well-designed tax system is that it is simple and economical to administer. It should result in the fewest possible disputes between the government and the taxpayers. The cost of collecting taxes should be low, and the cost to taxpayers of complying with the tax laws should be moderate.

Few individuals would argue against this characteristic of simplicity and economy. However, a tax system that yields as much revenue as ours does can hardly remain simple. The long legal battles fought over the definition of income indicate that some of our tax provisions are neither simple nor economical to administer, at least not when the cost to taxpayers of compliance is included. Large corporations, and even some wealthy individuals, must maintain a battery of accountants and lawyers so that they will not run afoul of the tax laws. On the other hand, our internal revenue department has demonstrated its efficiency by handling the mechanics of tax collecting at low cost.

3. A third characteristic of a well-designed tax system is that it should be fair. This characteristic naturally has stirred up much long and weary debate. Would the principle of fairness be served best by collecting the same amount of taxes from everyone since the rights of everyone as a

citizen are regarded as equal? Or would the interests of fairness be served better by assessing a rich man at a much higher rate than a poor man because the rich man has a greater tax-paying ability?

Possible principles of apportioning costs of government among taxpayers. Economists and taxing authorities long have given much attention to the criterion of fairness when considering the apportionment of tax burdens among taxpayers. Two general principles offer possible alternatives as basic policies to be followed in deciding how costs should be divided. We shall look at these principles separately.

1. The first alternative would be the *principle of benefits received.* If this were adopted, there would be an attempt to allocate tax burdens on the basis of the benefits that individuals received from government. A man who was greatly benefited by the services of government would pay higher taxes than a man who depended on government for few services.

This principle may seem fair at first glance, but very little thought is required to see how unworkable it would be. Under this principle a man in an insane asylum or a penitentiary who received his entire livelihood from government would be required to pay a high rate of taxes. On the other hand, a rich man who sent his children to a private school and who hired his own guards to protect his property would pay much less since he received fewer benefits than the man in the asylum or prison. The example, of course, is an exaggeration, but even under more ordinary circumstances the benefit principle does not work in practice for an entire tax system. A rich man's house may be fifty times as valuable as a poor man's cottage and the wealthy individual, therefore, may pay fifty times the amount of taxes paid by his poor neighbor. Both receive police protection and other services from government, but it is unlikely that the rich man gets fifty times as much service as the poor man.

The benefit principle may work very well in the case of fees assessed for certain special services. When a unit of government performs a service that directly benefits one individual and is only of minor benefit, if at all, to the rest of the community, the person so benefited should defray the cost. Special assessments for improvement of streets in front of a specific piece of property fall in this category.

2. A second alternative for apportioning the cost of government is the *principle of ability to pay.* Under this principle, the criterion of fairness is best achieved when the cost of maintaining government is borne by those members of the community most able to do so. In the United States

and in most other nations, the sums that must be collected from taxpayers are so vast that it is not practicable to adopt any other principle. Moreover, if it is not pushed to an extreme, this principle has an important element of fairness. However, everyone with few exceptions benefits from government in a democratic society and everyone should contribute something to its support. In an extreme form, the ability-to-pay principle might result in such high taxes on the able and well-to-do members of the community that the will of the most capable individuals to work or to risk their capital might be impaired. More will be said about incentives later in the chapter.

A number of versions of the ability-to-pay principle have been suggested. For example, it has been contended that the *principle of equal sacrifice* should determine the apportionment of taxes. This simply means that each member of the community should bear a share of the tax burden that will make his sacrifice equal to that of every other member of the community. Since a rich man has a greater tax-paying ability than a poor man, it is presumed that the wealthy man could pay more in taxes and still make no more sacrifice in foregone consumption than his less fortunate neighbor.

The measurement of tax-paying ability offers many difficulties. Income usually is regarded as the best yardstick, but it is widely recognized that two individuals with identical incomes may have very much different tax-paying ability. One, for example, may have more dependents than the other. The ownership of property also is not necessarily a good yardstick. The home of one individual may be heavily mortgaged, and amortization charges may be costly. Identical property owned by another person may be clear of all debt. The property of one man may consist of vacant lots; that of another may be equal in dollar value but may consist of a very profitable apartment building. Expenditure is another possible criterion, but like the others it is not always a just criterion of tax-paying ability. A large family may be forced to spend all its income to stay alive, while a small family of equal income may spend only a portion and save the remainder.

Despite these problems of measurement the ability-to-pay principle has become recognized as the fairer criterion, and income, although not a perfect yardstick, is regarded in most cases as a better measure of tax-paying ability than either property or expenditures. However, the benefit

principle should also be considered in apportioning some kinds of taxes. Moreover, it is good public policy to require everyone to contribute at least something to the support of government in return for the benefits that he receives. Everyone who pays taxes tends to have an interest in how the money is spent, and to that degree he is a more responsible citizen.

Formulation of tax rates. The ability-to-pay principle should be an important factor in selecting the type of rate schedule to be employed in a tax system. Before we define the various kinds of rate structures, let us examine the mechanics of rate determination. Before ascertaining how high or low a particular tax must be, we need to know two things: (1) the amount that we must raise by this tax and (2) the tax base, or what it is we are going to tax and how much there is of it. In the case of an income tax, the base is the incomes received by residents of the tax jurisdiction. In the case of a real-estate tax, the base is the value of taxable real estate located within the tax jurisdiction. With the amount we need to raise, and the base in mind, we can determine the tax rate. We merely divide the amount to be raised by the base. This can be expressed as follows:

$$\text{Tax rate} = \frac{\text{amount to be raised}}{\text{tax base}}$$

Several kinds of rate structures have been used or are being used now in tax systems. We shall examine the following four:

1. REGRESSIVE. If the tax rate decreased as the base upon which an individual is assessed increases, a tax is said to be regressive. In other words, a rich man pays a lower rate than a poor man. No tax authority would dare, in a democratic state, to levy an obviously regressive tax. However, a tax can become regressive in effect if it takes a larger percentage of income from the poor than from the rich. Moreover, any tax system that depends largely on consumption taxes and very little on income taxes may easily become regressive in that the poor pay out a larger percentage of their incomes than the wealthy.

2. PROPORTIONAL. A tax is proportional if an individual pays the same rate regardless of the change in the amount of the base. Most taxes on property are proportional. Under a proportional tax system, a man with a $10,000 piece of property would be taxed just twice as much as a man who had only a $5,000 piece of property.

3. PROGRESSIVE. A tax is progressive if the rate increases as the base increases. Most income taxes are progressive. Thus an individual with a large income not only pays a greater number of dollars than an individual with a small income; he also is taxed at a higher rate. The income tax systems in most nations now are steeply progressive. For example, the United States income tax in 1946 began at 20 per cent with the first $2,000 of taxable income and mounted to 91 per cent on all income over $200,000, less a 5 per cent reduction in net tax. A progressive tax usually is justified on the basis of ability to pay. Advocates of the progressive income tax maintain that ability to pay increases at a faster rate than income.

4. DEGRESSIVE. A tax is degressive if it is progressive up to a certain rate and then it becomes proportional. The United States income tax is degressive in the sense that the basic rate in 1946 on all income over $200,000 a year was 91 per cent. If progression is carried far enough, a tax automatically becomes degressive. In one foreign country during the Second World War, the excess profits tax finally was raised to 100 per cent and, one tax official sagely remarked, "That is about as high as we can put it." The tax had become degressive.

How high can tax rates go? The past few decades have been trying ones from the standpoint of government finance. The First World War with its enormous outlays by most major nations was followed after a few years of relative prosperity by one of the most severe depressions that the world has ever known. Most governments were forced to spend large sums for economic reconstruction. On the heels of the great depression, and perhaps related in part to it, came the Second World War, the most expensive war in the history of the world. The cost of the war to the United States alone was about 350 billion dollars, or roughly ten times the cost of the First World War. Costs directly related to the war lingered on after the end of the shooting in the form of demobilization, payments to veterans, and help to destitute foreign peoples. Our budgets, as well as those of other nations, have remained high and the burden on taxpayers has remained high. The question naturally arises, Are there limits to what can be collected from taxpayers? The answer is yes. We find that there are limits to what the tax collector can reap from the people, although some of the limits are quite flexible. Let us examine briefly three of them:

1. LEGAL LIMITS. In some tax jurisdictions, there are legal limits to the amount of taxes that can be levied or the rates that can be charged. Such limitations are imposed on state or local governments for the most part and can be changed either by the state legislature or by amending the state constitution.

2. POLITICAL LIMITS. In a democratic society the taxpayers can virtually place a limit on what they will pay in taxes by expressing disapproval at the polls when rates become what they regard as too high. In other words, there is a limit beyond which taxpayers will not go, and most officials are keen enough to know when the ceiling is being reached. Such limitations are somewhat relative, however. Taxpayers may object violently if rates are increased by large amounts, whereas little protest may result if identical rates are attained through a succession of small advances.

3. ECONOMIC LIMITS. There are certain limits above which it is not expedient to increase rates because of the effect that such high rates would have on taxpayer morale. In some cases, a low tax rate will yield more revenue than a high one because of the spur given to economic activity by low rates and the brake applied by high rates. The absolute economic limit to a tax rate would be that rate which took everything but a bare living from taxpayers. However, the practical limit would be reached long before such a crushing burden was imposed. Any rate that encroaches seriously on the taxpayers' accustomed standard of living obviously is near the economic limit.

These three limitations, although very real, are quite flexible. In the case of the political and economic ceilings, they automatically rise in times of national emergency such as a war, when it is regarded as patriotic to pay taxes. Moreover, the limit to tax-paying capacity depends in large measure on how the government spends the money it collects. The ceiling will be high if government expenditures are useful, productive, and flow quickly back into the income stream. On the other hand, the ceiling will be low if government expenditures are wasteful or add little to the level of incomes.

Question of incentives. Since we have seen that tax rates apparently must remain high if governments are to meet all their obligations, the question of incentives becomes of great importance. Tax rates must not be so high or so unfair that the will to work is hindered. Large incomes must not be so completely taken by the tax collector that the taxpayer

regards it as useless to expend the effort required to earn a large income. The starting of new business enterprises entails risk, and tax rates must not be so high that there is no opportunity for large profits even if an enterprise is highly successful. On the other hand, large sums must be raised in taxes and individuals with large incomes have greater tax-paying capacity as a rule than those with small incomes. The problem is not an easy one.

Some careful thinkers point out that the free-enterprise system depends on the efforts and willingness to venture of millions of individuals. They say that steeply progressive income taxes are a threat to the free-enterprise system because they take away the rewards of effort, initiative, and risk. Certainly, the tax system in a free-enterprise system like the American economy should be designed to encourage individuals to work, produce, and assume risks.

Other sincere students maintain that the problem of progressive taxes is not serious even in a free-enterprise system because of the importance of nonpecuniary incentives. It is argued that a person's work or his business has many of the elements of a sport or game. We find fun in our work or in running a business. Advocates of this view say that such nonfinancial rewards will keep us at work even if we have to pay high taxes.

These conflicting arguments are not easy to resolve. Certainly taxes could conceivably be raised high enough to destroy incentives. In fact it seems probable that tax rates as high as those that prevailed during the war years would seriously impair risk taking and effort in normal peace years. However, there is no reason to believe that the progressive income tax, or any other kind of progressive tax, will impair incentives if the rates remain within the bounds of what taxpayers regard as fair and reasonable. The principle of progression in itself has been accepted by taxpayers all over the world.

Taxes for nonfiscal purposes. So far we have discussed taxes in terms of the revenue that a government requires to meet its expenses. Taxes have been assessed many times in the past for nonfiscal purposes. For example, one Russian czar thought the cause of cleanliness should be encouraged so he placed a tax on beards. Even now some nations seek to encourage matrimony by levying special taxes on bachelors. In many nations tariff duties are assessed, not with the objective of raising revenue, but to keep foreign goods from being imported. Highly progressive income tax rates sometimes are encouraged on the grounds that they

equalize incomes. This result is regarded by some as desirable. High death duties also are sometimes justified on the grounds that it is in the public interest to break up large estates rather than to let them grow as they are passed down from generation to generation.

When taxes are levied for nonfiscal purposes, the results are not always desirable. In the case of protective tariffs, some industries benefit by the reduced competition from foreign producers. On the other hand, consumers may have to pay higher prices for domestic products. Similar controversies might be cited in other cases where taxes are levied for nonrevenue purposes. Some such taxes may, indeed, be justified, but they should be levied only when there is general agreement that the end sought is a desirable one.

Kinds of taxes. Before we examine the tax systems of our national, state, and local governments, we shall consider briefly some of the different kinds of taxes that have been used to produce revenue for governments. Any student of taxation is struck by the ingenuity of man in the multitude and variety of taxes that he has been able to devise. Almost everything we eat, wear, use, own, or earn bears a share of the tax burden whether we realize it or not. In some cases we know that we are being taxed because we get a bill from the tax collector. In other cases we do not know that we are paying a tax because it is included in the purchase price of something we buy. Certain kinds of taxes have come to be of special importance. Let us examine several that are bulwarks of our revenue system.

1. Income Taxes. In this type of tax, the base is income. Most income taxes in the United States, as well as other countries, are net income taxes. This means that certain deductions and allowances from gross or total income are permitted before the tax is computed. The cost of producing income usually is deductible, but the question of what constitutes cost often results in disputes between the taxpayer and the revenue collector. Moreover, certain kinds of income are not taxable. Before tax rates are applied, net income must be established. If the income tax is progressive, the income is divided into brackets and higher rates apply for each bracket up the scale.

As we already have indicated, income is one of the best yardsticks of tax-paying ability. Correlation, however, is not perfect because different taxpayers have different personal problems. In the long run all tax revenue has to come from someone's income. From the standpoint of

the government, heavy dependence on income taxes may mean wide fluctuations in revenues. We have pointed out, however, that some students of taxation would regard wide fluctuations in tax collections from boom to depression as desirable. This would be true only if a government had almost unlimited borrowing power at its command.

2. Excise Taxes. Excises are taxes levied on the value of a domestic commodity or at a certain specified amount per unit of the commodity. They are rather painless to the taxpayer in the sense that he does not know he is paying them because the tax is included in the purchase price. Many of the commodities that we buy bear excise taxes. Among the articles of mass consumption heavily taxed in this manner are liquor, gasoline, and tobacco. In some cases, we do realize that we are paying such taxes because that fact is pointed out to us. For example, on every gasoline pump the price per gallon is broken down into the price of the gasoline and the amount of the tax. Few motorists, however, appear to give much thought to the matter. They are interested in the net cost of the gasoline. If they thought the matter over carefully, they might be more insistent on having gasoline taxes go entirely for highways and other motoring aids, and they might object to having them siphoned off for other governmental purposes.

Excise taxes sometimes are criticized on the grounds that they make the poor man pay out a larger percentage of his income than the rich man. This contention is made even though every purchaser of a particular unit of a good pays the same amount of tax. It should not be overlooked, however, that different commodities bear excise taxes at different rates, and that luxuries often are taxed at higher rates than necessities. Excises often are defended on the grounds that they require contribution to the upkeep of government by persons who might otherwise escape all taxation.

3. Tariffs. Tariffs are a form of tax collected when goods are imported. As we pointed out in Chap. 19 on International Trade, tariffs designed to produce revenue would be low enough not to impede the importation of goods. Some commodities can stand higher tariff rates than others, however, and rates in the tariff schedules of any nation usually vary over a wide range. Protective tariffs are tariffs that are purposely set so high that foreign goods cannot easily be imported and sold in competition with domestic goods.

4. Sales Taxes. Most United States sales taxes are levied by the states and are retail sales taxes. Thus they are taxes levied at the time of final sale and are based on the retail selling price. If commodities are taxed both at the wholesale and retail level, we use the term "general" sales tax. In some foreign countries there are manufacturers' sales taxes, levied when goods first pass from the manufacturer to the distributor.

Sales taxes usually apply to tangible property. However, a gross receipts tax includes all the items covered by a general sales tax as well as income from personal and professional services.

A sales tax is similar to an excise tax in the sense that it is levied on a commodity offered for sale. Excise taxes, however, often are levied on a per unit basis instead of a price basis, whereas a sales tax is always a percentage of the price. Moreover, the rates of excises are different for different commodities, and the amount of the tax is incorporated in the purchase price by the time the article reaches the consumer. On the other hand, a sales tax usually is assessed at a uniform rate for all commodities and is pyramided on top of the selling price. Sales taxes often are criticized on the grounds that they take a higher percentage of income from the poor than from the wealthy. They are defended because they are dependable revenue producers. They also are defended on the grounds that they relate the payment of taxes to the amount of spending. An individual who spends recklessly has to pay more in taxes than a frugal person.

5. Property Taxes. The property tax, bulwark of local governments, is a tax on wealth. A general property tax would be a tax on all wealth, tangible and intangible, including both real estate and personal property. Personal property consists of such tangible items as merchandise, furniture, automobiles, and books; and also such intangibles as mortgages, stocks, bonds, money, and bank accounts. In a number of areas personal property is not taxed, and the property tax thus becomes solely a real-estate tax levied on land and on such improvements on land as industrial buildings, commercial buildings, and residences.

The property tax has been widely criticized on many grounds. Some contend that property is a poor measure of tax-paying ability since much property, such as a home or vacant lot, does not produce income. Moreover, it is contended that the property tax does not reach individuals who, although they own no property, possess much tax-paying ability because of skills and competence to render highly remunerative serv-

ices. In addition, it often has been demonstrated that heavy dependence on property taxes may result in inequality of the tax burden which must be borne by individuals residing in different tax districts. This is due to the unequal distribution of valuable property as between different taxing districts. It also is due to different assessment practices.

OTHER TAXES. Many other kinds of taxes are collected in the United States and elsewhere. Some of these will be mentioned in our discussion of state and local tax systems. Among important sources of revenue in some localities are taxes on motor vehicles, and severance taxes on natural resources such as the tax in many states and nations on petroleum when it flows from the wells. Other taxes, which sometimes yield substantial revenue, include special license fees, various types of business taxes, taxes on admissions, and special assessments. The tax collector has no dearth of things to tax.

Fiscal systems of national, state, and local governments. The Federal government has been relying to an increasing extent on income taxes. Personal income taxes are levied on the incomes of individuals. Incomes of corporations are taxed at a different scale of rates. At the present time about two-thirds of total Federal tax collections come from income taxes. Just before the Second World War only 35 to 40 per cent of total revenues was from this source.

Excise taxes and custom duties were long the mainstays of the Federal fiscal system, but they have been decreasing in importance. Before the First World War they contributed 85 per cent of total tax revenue. Excises alone contributed 40 per cent. Even in the late 1930's they contributed about one-third. Now they yield only about 15 per cent. There are numerous other sources of Federal revenues, for example, estate, gift, and employment taxes.

State governments have devised an imposing list of taxes in an attempt to swing away from dependence on property taxes, which formerly were an important source of revenue. More and more states are leaving the property tax for the support of local governments, although this tax still provides substantial sums for some states. Among the important sources of income for states are sales taxes, income taxes, inheritance taxes, and motor-vehicle taxes. Taxes on alcoholic beverages and tobacco are also important, and business license fees are income producers in some states. Grants from the Federal government are still another source of income.

A study of the sources of income for different states indicates a wide divergence in major sources of revenue. Some states rely heavily on a particular tax. As an example, Illinois obtained nearly 40 per cent of its total income from the retailers' occupation tax in a recent year. Some states find that severance taxes, levied on natural resources such as metals and petroleum, are important. For example, petroleum flowing from Texas oil wells has provided large revenues for that state.

We can classify the income of local governments, which include counties, cities, towns, and all other local units, into four groups: (1) taxes, (2) aid received from the state or Federal government, (3) charges for services, and (4) miscellaneous income. The bulwark of income for local governments is the property tax. This tax provides around 70 per cent of the total income of local governments and in excess of 85 per cent of all tax revenues. Some units of local government receive nearly all their income from this tax.

License fees and permits are another source of tax income, particularly for municipal governments. Some cities have levied sales taxes and several, including Philadelphia, have income taxes. Amusement taxes, taxes on hotel rooms, cigarette taxes, taxes on taxicabs, alcoholic beverages taxes and many others have been levied by hard-pressed cities where property tax revenue has proved inadequate.

Grants from other layers of government and shares of taxes collected by higher layers of government have been consequential for local governments. Grants may be made from a state to a county or city to aid schools. Grants may be made for upkeep of arterial streets traversed by state highways. Local governments have received aid from higher units for public welfare.

Local governments also receive income from the charges they make for various services. Examples are recording of documents and inspection fees. Fines also provide income. Recently, charges have been made by some hard-pressed cities for such essential services as garbage collection. Cities also have many miscellaneous sources of revenue. Some, for example, have recently found that parking meters provided needed income. Others own various municipal enterprises such as golf courses. Others own their own public utilities which sometimes provide income for other governmental functions. Some receive a portion of the gross revenues of utilities.

Shifting and incidence of taxes. Frequently the individual or business firm on whom a tax is levied is not the one who ultimately pays it. The tax may be passed on to someone else in the form of a higher price charged for a product. In this case we can say that the tax has been shifted. The incidence of a tax is its final resting place. Thus, if a tax is levied on a food manufacturer who passes it along to the consumer in the form of a higher price, we say the incidence is on the consumer. This is true despite the fact that the government levying the tax may have collected it from the manufacturer.

Federal debt. The national debt of the United States has been pushed up after each of the five major wars in which the nation has engaged. Each time up to the Second World War the debt has been nearly paid off or very sharply reduced. However, the huge debt created by that war may not be so easy to reduce. In June, 1940, our gross public debt was 43 billion dollars. In early 1946 it was approximately 279 billion. This rise was directly related to the Second World War, during which only 46 per cent of the required revenue was raised by taxes, the remainder coming from borrowing. Despite the fact that less than half of the war was financed through tax collections, our showing in this regard was better than in the preceding war when only 28 per cent of the cost was met by taxes.

The fact that our public debt has become so large raises questions that are not easy to answer. We might ask, How can we pay back the debt with the least amount of burden on the taxpayers? How rapidly should we try to repay it? Should we even try to reduce it? Who should hold the bonds—individuals, banks, insurance companies, etc.—if we are to carry the huge debt with the least amount of burden? Much difference of opinion usually develops when such questions are raised. Obviously, to extinguish the debt would take many years under any conceivable schedule of repayment. Even if we could pay 5 billion dollars a year, more than half a century would be needed. Moreover, we can be quite sure that the U. S. government will not operate in the black every year. The black figures in the fiscal year 1947 were the first since 1930. In every intervening year, the government spent more than it took in and the debt climbed. It seems more probable that, if we repay the debt at all, a long period of time will be required.

As a matter of fact, the history of public debts shows that they rarely are paid off. Our experience has been better than that of most nations.

Although Great Britain long was regarded as a pillar of financial rectitude, the British have done no better than whittle away at their debt between wars or periods of financial strain. Their debt from the Napoleonic wars, for example, was never paid off or even sharply reduced. In Great Britain and in other solvent nations, individual bondholders get their money back when their issues fall due, but the money is obtained from other lenders. Unfortunately, there have been many examples in the history of public debts where governments have repudiated their obligations.

During a period of rapid growth, a nation often finds its public debt to be a decreasing burden even though the amount remains the same or rises at a lesser rate than population, income, and wealth. This was the experience of the British during the nineteenth century. The problem becomes more serious when the increase of the debt outstrips a nation's wealth and income.

The argument often is advanced that the size of a debt makes no difference when it is held internally. Our Federal debt is an internal one since almost all United States securities are owned by our citizens. Those who make such a contention say that when a debt is internally held we owe it to ourselves. They argue that we merely collect taxes from our people and pay the money back to them when we disburse interest on the debt or redeem bonds. This is compared to taking money out of one pocket and putting it in another.

This point of view neglects the fact that we rarely collect taxes from individuals, banks, insurance companies, and others in the same ratio that they hold government securities. Taxes collected for debt service from low-income groups often increase the already unequal distribution of wealth, since most of the bonds are owned by upper income groups, or by banks and other institutions. In addition, the strains, threat to incentives, and annoyances resulting from high taxes for debt service may be quite important.

Summary

In this chapter we have noted that the cost of government has been rising. A big share of the responsibility was attributed to the costly wars in which the nation has engaged. Another very important reason, however, is the increase in the services that the citizens are demanding from nearly all layers of government. We have indicated also some of the fundamental principles of

taxation, such as rate determination, allocation of tax burden, and the characteristics of a good tax system. We have explained the principal kinds of taxes and noted the kinds of taxes that have proved to be important revenue producers for various layers of government. We have also pointed out that the individual upon whom a tax is levied does not always bear the burden. Frequently, the tax is shifted to someone else. Finally, we noted the rise in the public debt and the relation of this rise to the Second World War.

Study Questions

1. Why has the cost of government risen in the United States?
2. What are the two principal sources of government revenue?
3. Define a tax.
4. What is the distinction between a tax and a fee?
5. What are the characteristics of a well-designed tax system?
6. State several principles that have been proposed for apportioning tax burdens among individuals.
7. What do we mean by a tax rate? How are tax rates formulated?
8. Define each of the three types of tax-rate structures described in this chapter.
9. What are some of the factors that tend to keep tax rates from rising beyond certain limits?
10. Why does the question of incentives become important in considering tax rates?
11. What are some of the possible objectives when taxes are levied for nonfiscal purposes?
12. What do we mean by a sales tax? What are some of the different kinds of sales taxes?
13. What do we mean by a net income tax?
14. Distinguish between an excise tax, a sales tax, and a tariff.
15. Define a general property tax.
16. Summarize the principal sources of revenue for national, state, and local governments.

Exercises

1. Assume that a certain government depends entirely on the property tax. Assessed valuation in its jurisdiction totals $55,000,000. The local government needs to raise $4,500,000. What must the tax rate be?

2. Suppose that the following distribution of average taxable income, average taxable property, and average annual expenditure prevails in a community:

Percentage, families	*Average taxable income*	*Average taxable property*	*Average annual expenditures*
48	$ 2,000	$ 500	$1,900
32	4,000	3,500	3,600
12	6,000	7,000	5,200
8	10,000	9,000	7,500

The income tax rates in effect are as follows:

0% on first	$2,000
20% on next	2,000
35% on next	2,000
42% on next	4,000

Property is taxed at a rate of $3 per $100 of property. The community has a 2 per cent general sales tax which covers all consumer expenditures. Is the tax system of this community regressive, proportional, or progressive? Prove your answer.

3. Of the kinds of taxes described in the chapter, which do you regard as most desirable in a country where (*a*) great inequality of incomes and wealth prevails, (*b*) incomes and wealth are distributed about equally? Have arguments ready to defend your answer.

References for Further Reading

Blodgett, R. H.: "Principles of Economics," Rinehart & Company, New York, 1946, Chap. 28.

Bye, R. T., and W. W. Hewett: "Applied Economics," F. S. Crofts & Co., New York, 1941, Chaps. 23 and 24.

FAIRCHILD, F. R., E. S. FURNISS, and N. S. BUCK: "Elementary Economics," The Macmillan Company, New York, 1939, Vol. II, Chaps. 29-32.

GARVER, F. B., and A. H. HANSEN: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 32.

ISE, JOHN: "Economics," Harper & Brothers, New York, 1946, Chap. 33.

26. Labor Organization

Organized labor is an integral part of the modern free-enterprise system, no less so than the factory, the corporate form of business organization, or financial institutions; and today organized labor plays a more important part in the economic process than ever before. The purpose of the present chapter is to give a brief description of the labor movement in the United States and to show how it has modified the way in which our economy operates.

Development of labor organization in the United States. The first attempts of labor to organize in the United States occurred late in the eighteenth century in certain of the skilled trades. In some cases the workers organized solely to carry on a strike and disbanded as soon as the strike was over; in other cases, however, workers organized trade-unions on a more permanent basis.

The members of these early unions were handicraft tradesmen like shipwrights, carpenters, shoemakers, and tailors. They had not yet been affected by the industrial revolution and the machine. However, as the centers of trade and industry grew, these workmen came in contact with one another in constantly increasing numbers, and they soon saw the possibilities of improving their situation by joint action.

Once the labor movement had become established in this country, it continued to grow but in a rather sporadic fashion. In periods of prosperity new unions would be established and their membership and influence would increase, but in periods of depression membership would decline and sometimes the unions themselves would disappear. Nevertheless, as the nineteenth century progressed, labor became an increasingly important factor in the economic life of the country. For the most part, however, the success of the labor movement was confined to workers in the skilled trades. Even well into the twentieth century after the United States had become the greatest industrial country in the world, the vast majority of workers in our factories and mines were still

unorganized. Only in very recent years have the workers in our great mass-production industries, like steel and automobiles, been brought into the ranks of organized labor.

The first permanent trade-unions on a national scale were formed in the decade just preceding the Civil War. However, attempts to achieve national unity of the labor movement as a whole had no marked degree of success until after the formation of the Knights of Labor. This organization attempted to bring together all types of unions and all types of workers, and in 1886 at the height of its success it had reached a membership of over 700,000. Many craft unions were affiliated with it, but it also included general labor unions which admitted to membership all workers regardless of occupation or skill. For a short time its power and prestige were great, but after 1886 it declined very rapidly. One of its weaknesses was the attempt to bring together in the same unions both skilled and unskilled workers, as well as workers from a variety of industries. Such groups did not have enough interests in common to enable them to survive.

The American Federation of Labor, organized in 1886, achieved a much more permanent success than the Knights of Labor. The AFL is primarily an association of trade or craft unions. After the decline of the Knights of Labor it became, and for many years remained, our one great national labor federation; but today it has a strong rival in the Congress of Industrial Organizations, or CIO.

The CIO, in its early stages known as the Committee for Industrial Organization, grew out of a dispute at the 1935 convention of the AFL. Many union leaders, including John L. Lewis of the United Mine Workers, felt that the mass-production industries, like steel and automobiles, could be organized only by bringing all the workers in each industry, regardless of the kind of work performed, into a single union. Such an organization is called an *industrial union,* and it differs radically in principle from the *craft union,* which is open only to the members of a certain skilled trade, for example, carpenters or machinists. However, the majority of the unions belonging to the AFL were organized on the craft principle, and they were opposed to organizing the mass-production industries on an industrial basis. They wanted to bring the skilled workers in these industries into their own ranks.

A strong minority at the AFL convention insisted that organization of the mass-production industries on a craft basis was wholly impractical,

and they urged that the AFL undertake a vigorous campaign to organize these industries on an industrial basis. Since the majority opposed this bitterly, a long controversy resulted; and after the convention the representatives of an important group of national unions formed the Committee for Industrial Organization. Later this group, under the leadership of John L. Lewis, broke off all relations with the AFL and organized on a permanent basis as the Congress of Industrial Organizations. The CIO has had astonishing success in creating strong industrial unions, especially in steel and automobiles. Today it rivals the parent AFL in membership, and both groups have grown tremendously.

In addition to the unions affiliated with either the AFL or the CIO there are a number of important independent unions. Most notable of these are the "big four" railroad brotherhoods—the engineers, the firemen, the conductors, and the trainmen.

As has been noted, until recent years organized labor has included a very small minority of American workers. At the low point of the depression, in 1933, the total membership of all labor unions in this country was estimated at less than 3 million. However, business recovery and favorable New Deal legislation soon brought a great increase, so that by 1937 it was estimated that total union membership exceeded 7 million, but the Second World War and the early postwar prosperity proved an even greater stimulus. As a result, organized labor reached an all-time peak in both power and numbers with a probable union membership of close to 15 million.

Reasons for the development and growth of labor unions. Before the industrial revolution, production was carried on by handicraft methods in small home shops, and the typical handicraft worker was an independent businessman. There were, to be sure, apprentices who were learning the trade and also some journeymen who worked for wages, but most of these apprentices and journeymen expected sooner or later to go into business for themselves.

The industrial revolution brought a great change in the situation. In order to use power-operated machinery, it was necessary to build a factory, and a factory represented a large investment and required a great number of workers. The result was that the new factories came to be owned by a relatively small group of wealthy capitalists, while the ordinary worker ceased to be a businessman and became merely a wage earner. Gradually the distinction between the owners of industry on the

one hand and the workers on the other became more firmly established; and as the typical factory or mill increased in size, the individual worker became increasingly aware of his relative unimportance and helplessness. For the worker the remedy for this situation lay in organization. By uniting with his fellow workers he could deal with his employer on something like equal terms, and thereby achieve a greater degree of security and self-respect.

The effect that organization had on the morale of the worker and on his whole attitude toward life is often overlooked or minimized by those who study the labor movement from the outside or from the viewpoint of the employer; but it has been testified to by workers time after time in the most convincing terms. Many a coal miner, for example, has described the difference between a nonunion and a union mining community as like the difference between black and white. The unionized workers have a sense of confidence and self-respect which they say transforms the entire life of the community.

Among the particular objectives that induce workers to organize, higher wages and shorter hours take a prominent place. Another important objective is better working conditions. Working conditions cover a great variety of specific matters. In addition to provisions for health and sanitation they include such things as discipline, facilities for eating lunch, and opportunities for rest and recreation. Still another factor of importance in inducing workers to organize is security. Included under security are protection against being dismissed for no good reason; seniority rights that give preference for employment or for promotion; and pension and disability rights. Last, but not always least as an incentive for unionization, is the possibility of bringing political pressure and obtaining the passage of legislation favorable to workers.

One factor that should not be overlooked in discussing the reasons for the growth of the labor movement is the influence of the spread of free public school education. Before the nineteenth century the common workingman had little influence either on government or his own position in society; and one of the principal reasons for this was his almost complete ignorance of everything outside of his own narrow experience. When workers were taught to read and write, and when later some even went to high school and college, all this was changed. For the first time it became possible for large numbers of workers to understand their economic position, to communicate freely with one another, and to

organize to protect their common interests. Organized labor has recognized the advantages that public schools confer on the workingman, and it has generally given strong support to efforts to extend free schooling to the college and university level.

Structure of organized labor. Although unions vary considerably in structure and the way in which they are controlled, most of them have certain features in common. The basic structural unit of either a craft or an industrial union is the *local.* In the case of smaller communities, the local may consist of all the workers in a craft or industry in a given town or county. In large cities, however, such an organization would be unwieldy; and therefore separate locals are formed among the workers living in a certain section of the city or among those employed in a certain plant. Where a union has a number of locals in the same area, cooperation is secured by the formation of a *district council.*

Although the local is the basic structural unit of organized labor, the center of power is usually found in the *national union.*[1] Examples of national unions are the Brotherhood of Locomotive Engineers, the United Mine Workers, and the International Ladies Garment Workers. Originally, national unions were only loose federations, but as the advantages of united action became apparent, the locals gradually relinquished to them more and more power. Today, as a rule, the national union makes the rules and regulations by which the locals must abide. The national union also controls strike, disability, and other benefit funds, and these constitute its most effective means of controlling the locals. A threat to withhold benefits is usually most effective in inducing locals to submit to the authority of the national.

A *general federation* is a combination of different craft or industrial unions for the purpose of furthering their common interests. Federations may be either local or national. A local federation is often called a *city central.* The two great national federations have already been mentioned, the CIO and the AFL. Federations are voluntary associations which have little control over their member unions. They cannot call strikes, and their influence and prestige depend chiefly on the voluntary support of the national unions that are affiliated with them. The member unions of either the CIO or the AFL are free to withdraw whenever they please.

[1] Sometimes a national union is called an *international* because it includes locals in Canada.

Collective bargaining. Most of the objectives that unions seek can be obtained only through dealing with employers. When the representatives of an organized group of workers meet with an employer and attempt to secure an agreement concerning hours, wages, or other conditions of work, the resulting negotiations are called *collective bargaining.*

If an employer and the representatives of a union cannot reach a direct agreement on the points at issue in a dispute, they may call in some prominent and impartial citizen to give them advice and counsel. This method of seeking a settlement is called *mediation.* If they still find it impossible to reach a settlement, they may then agree to submit their dispute to *arbitration.* In this case they select a mutually satisfactory arbitrator and agree to abide by his decision; or they may set up a special board of arbitration or call in a government board already established.

Once the points at issue have been settled, a written agreement is drawn up specifying the obligations of both employer and employees. Usually such an agreement, or contract, runs for a definite period of time, and often it is subject to renegotiation if either party gives a certain number of days' notice.

Weapons of organized labor. The principal weapons of organized labor are the strike and the boycott. A strike is a mass stoppage of work, while a boycott is a mass refusal to trade with an employer. In the early days of the labor movement courts frequently held that it was illegal for workers to organize for the purpose of carrying on strikes or boycotts. Strikes and boycotts were held to be in restraint of trade, and therefore they were illegal under the common law. Furthermore, under the common law, for a group of people to cooperate in carrying out an illegal purpose constituted criminal conspiracy. Today, however, the right of workers to organize and to enforce their demands by peaceful strikes or boycotts is firmly established, provided their objectives are such "legitimate" ones as higher wages, shorter hours, or better working conditions.

Generally the courts make a distinction between *primary* and *secondary* boycotts. Primary boycotts involve only a refusal of the workers themselves to buy from an employer with whom they have a dispute. Secondary boycotts involve third persons who are not parties to the disagreement in question. In a secondary boycott the union, for example, boycotts other firms who sell the goods of an employer involved in a labor dispute or who carry on business relations of any kind with him.

A union often attempts to increase the effectiveness of both strikes and boycotts by *picketing,* that is, by having union members stand outside a place of business to urge nonstriking workers to join the strike or to urge customers not to buy goods. Picketing is generally legal when it is peaceful and when it does not involve keeping workers or customers out of an establishment by force or threats.

In rare cases, organized labor has resorted to *sabotage,* that is, to the destruction of machinery or materials in ways that are difficult to discover. This represents one of the most vicious types of industrial warfare and exposes the perpetrators of the acts, if caught and convicted, to the full penalties of the law.

The company union versus the independent union. A *company union* is a union limited to the workers for a given firm. Usually the officers of a company union are elected from among the employees themselves and devote only part of their time to work for the union. An *independent union,* unlike a company union, is controlled entirely by the workers themselves and not by the employer. In negotiations it is represented by officials who work full time for the union and who therefore are not dependent on the employer for their job.

Organized labor generally is opposed to company unions. It believes that only an independent union can really protect the rights of the workers.

The closed versus the open shop. A *closed shop* is a plant in which, by agreement between the union and the company, only union men are employed. As a rule unemployed union men are given preference for jobs, but in case a nonunion man should be hired he must join the union at once. An *open shop* is a plant that employs both union and nonunion men. Unions naturally seek to obtain a closed-shop agreement, because such an arrangement increases their prestige, their financial strength, and their bargaining power. Employers, on the other hand, are likely to oppose the closed shop because they wish to keep the power of organized labor within limits. Many people who have no direct interest in labor controversies also oppose the closed shop because they believe that requiring a man to join a union in order to get a job is an infringement of personal liberties.

The closed shop is often combined with the *checkoff.* Under this arrangement the company automatically deducts the union dues of each worker from his wages, and turns the sums deducted over to the union.

The closed shop and the checkoff combined contribute substantially to what is sometimes called "union security."

The *union shop* is a compromise between the closed shop and the open shop. Under this plan nonunion workers may be employed, but such workers must join the union within a specified period of time.

Antiunion activities of employers. In their attempts to deal with unions or to prevent organization of their workers, employers have resorted to various devices. The *lockout* corresponds to the strike. Instead of the employees stopping work in order to enforce their demands, the employer shuts down the plant in order to enforce his demands. The *black list* is a list of the employees who have been obnoxious to the employer because of their prominence in union activities. Such a list is passed from one employer to another, so that those on it have difficulty in getting jobs anywhere in the industry. The *yellow-dog contract* is an agreement signed by an employee when he takes a job to the effect that he does not belong to a union and will not join one during the period of his employment. Other methods used sometimes in the past by employers to fight unions include discharging active union workers, bringing in strike-breakers, and obtaining court injunctions against strikes, boycotts, or other union activities.

Norris–La Guardia Act. Organized labor has always resented the use, by employers, of court injunctions against what labor regards as legitimate weapons for enforcing its demands. It has alleged that the courts have, as a rule, been prejudiced in favor of employers. Sometimes an employer would succeed in getting a temporary injunction against a strike or boycott, and hearings on the injunction would drag out for weeks or months. Meanwhile the union involved would be powerless to act at a strategic time.

The Norris–La Guardia Act was passed in 1932 just as the great depression was reaching its worst stage, and it was generally regarded as a great victory of labor. It has substantially reduced the use of the injunction in labor disputes. Its most important provisions are in Section 4, which lists a number of acts of workers that Federal courts may not prohibit by injunction. These acts are as follows:

1. Stopping work.
2. Joining a union, regardless of any yellow-dog contract.
3. Paying benefits to strikers.

4. Aiding strikers to defend themselves in court.
5. Publicizing, without fraud or violence, the facts of a labor dispute.
6. Assembling peaceably to promote labor interests in a dispute.
7. Notifying any person of the intention to do any of the things mentioned.
8. Agreeing with others to do any of these things.
9. Inducing others to engage in the above acts, even though they have promised not to.

Labor under the New Deal. No president was ever more in sympathy with labor than Franklin D. Roosevelt, and under his administration organized labor increased greatly in numbers and in economic and political power. A good part of the advances made by labor under the New Deal can be attributed to a single piece of legislation, the National Labor Relations Act, often referred to as the Wagner Act. Partisans of labor frequently call this act the Magna Charta of the labor movement.

The National Labor Relations Act was passed in July, 1935. It was designed to protect workers in their right to organize, and also to encourage collective bargaining between unions and employers. Employers were prohibited from engaging in certain "unfair" practices. These forbidden practices included (1) interfering with the rights of employees to form independent unions, or to bargain collectively, (2) discriminating against employees for union membership or union activities, (3) controlling or supporting a union, financially or otherwise, and (4) refusing to bargain collectively. In addition the act provided for a National Labor Relations Board to administer its provisions. The board had power to investigate disputes, to settle disputes, to issue orders to employers, and to hold elections to determine which of two or more rival unions should be recognized as the bargaining agent for the workers in a given plant. If necessary, the board could appeal to the Federal courts to enforce its decisions.

Another piece of New Deal labor legislation of considerable importance is the Fair Labor Standards Act of 1938, otherwise known as the Wages and Hours Law. The purpose of this statute is to protect unorganized low-income workers by setting minimum wages and maximum hours of work. If workers are employed for more than the maximum number of hours, they must be paid time and a half for overtime. At present the

minimum wage is 40 cents per hour and the maximum number of hours that may be worked in any one week without payment of overtime is 40.

Labor in wartime and in the postwar period. Reference has already been made to the great advances of organized labor under the New Deal. During the war and the years immediately following, labor continued to forge ahead. This was partly a result of the favorable legislation already mentioned, and partly a result of the high level of prosperity and the great demand for workers.

However, at the same time that organized labor was experiencing this increase in power and prestige, there was a gradual but marked growth of antilabor sentiment among important segments of the general public. One cause of this was wartime strikes. Though such strikes did not set production back very greatly, some strikes did occur; and quite naturally they were deeply resented by the public at a time when pushing production to its maximum might save the lives of men on the field of battle.

Other factors contributed to the growth of antilabor sentiment both during and after the war. Many people who had formerly been in sympathy with the unions began to feel that organized labor was becoming a powerful minority pressure group and that the time had come when it should be subjected to restraints. Many also, rightly or wrongly, resented the tactics of certain widely publicized labor leaders. To literally millions, for example, names like John L. Lewis and James V. Petrillo became symbols of arbitrary power exercised in disregard of the public interest. In the period following the war, the popularity of organized labor was further reduced by several major strikes which threatened to paralyze all industry throughout the country. Most serious of these were nation-wide strikes in the steel and soft coal industries. Still another factor contributing to antilabor sentiment was the belief in some quarters that union wage demands were excessive and that rising labor costs were a principal cause of the rising cost of living.

Many of the people who felt that organized labor was gaining a dangerous amount of power were especially disturbed by the development of so-called "national labor monopolies." They used this term to apply to those national unions which had become so strong that they could tie up, not just one plant or one firm, but an entire industry. As a result of New Deal legislation and the war, several unions had grown to the point where they possessed this power. The most notable examples

were to be found in the automobile, steel, and coal industries. Those who felt that this development was a dangerous one urged that the principle of the antitrust laws should be applied to labor. They proposed that labor organizations should be limited to the employees of a particular firm, and that industry-wide collective bargaining should be outlawed.

Employers had criticized the Wagner Act on a number of grounds. They pointed out, for one thing, that although it forbade various unfair practices on the part of employers, it put no similar restraints on certain unfair practices sometimes engaged in by unions. Further, they said, it did not make unions responsible for their actions. If a union broke an agreement that an employer had accepted in good faith, the employer had no effective recourse. Also, they maintained that the act as actually administered tended to deprive an employer of his right of free speech. If he circulated pamphlets among his workers expressing his views on labor organization, this action was likely to be treated by the NLRB as an interference with the rights of his workers to organize independently.

Whether or not the Wagner Act swung the balance of power too far in the direction of labor for the national good is a highly controversial question. Certainly many people believed that it did, and therefore a strong demand developed for action to correct this situation. The result of this demand was the Taft-Hartley Act. The act as finally passed did not outlaw industry-wide collective bargaining, but it did place a number of important restraints upon the privileges and powers of organized labor.

Taft-Hartley Act. The Taft-Hartley Act was passed by Congress in June, 1947. It was passed over President Truman's veto and also over the violent protests of organized labor. It amended the Wagner Act in a number of important respects.

Some of the more important provisions of the Taft-Hartley Act are as follows:

1. It forbids unions to engage in certain unfair practices. Among these are jurisdictional strikes, secondary boycotts, excessive initiation fees, and forcing employers to pay for work not performed and not intended to be performed. This last prohibition is aimed at "feather-bedding" practices, like the requirement of the musicians' union that a qualified musician be present when phonograph records are played at a radio station.

2. It attempts to protect the employer's right of free speech.

3. It specifies that the obligation of employers and unions to bargain collectively shall not be construed to mean that either party must make a concession.

4. It provides that unions to be recognized by the NLRB must file with the Department of Labor reports on their finances and internal union practices.

5. It forbids the closed shop, and permits a union shop only if a majority of employees vote for it. Under the union shop arrangement, nonunion workers may be hired, but they must join the union within 30 days.

6. It permits the NLRB, but not employers, to seek court injunctions against jurisdictional strikes and secondary boycotts; although employers themselves cannot ask for injunctions, they are permitted to sue unions for damages in the Federal courts.

7. In emergencies where a strike would endanger the national safety, it permits the President to appoint a fact-finding board and ask the Attorney General to seek an injunction against the strike. If no settlement is reached within 80 days, the injunction must be withdrawn and the President can then seek aid from Congress.

8. The dues checkoff is forbidden except with the consent of individual workers.

9. Employers are permitted to contribute to union welfare funds only if they participate in the administration of the funds.

10. Unions are forbidden to make political contributions in connection with the election of the President or members of Congress.

There is not space here to discuss the Taft-Hartley Act in any detail. On their face a number of its provisions seem to be desirable checks on possible abuses of power by unions and union officials. Other provisions, like the prohibition of the automatic checkoff in a union shop, are debatable. The provision that employers can sue unions for damages resulting from illegal union activities seems a reasonable protection of the rights of employers. However, such a provision may prove to be subject to abuse. Conceivably, an employer could bring suits simply to harass unions and exhaust their funds in litigation. Though it is clear enough that the Taft-Hartley Act places substantial restrictions upon union activities, a considerable period of time will have to elapse before the nature and effect of these restrictions can be appraised with any

accuracy. A great deal will depend on the way in which the act is interpreted by the courts.

Economic effects of the growth of organized labor. The total effects of organized labor on our economy have been very great. Some of them are rather obvious, but others are less easily seen and have frequently been overlooked or misunderstood. One of the more obvious results of the activities of organized labor is the passage of much legislation favorable to workers. This type of legislation covers a wide field and includes such matters as workmen's compensation for industrial accidents, health and safety standards in factories, child labor, social security, hours of work, minimum wages, and the rights and privileges of unions.

In many other areas of our economic life the effect of organized labor is not so clear. Often unions have greatly complicated the problems of employers, especially those of small employers. A strong union, for example, may hold the whip hand over a retailer because it can completely cut off the delivery to him of goods. Occasionally, however, employers have testified that dealing with a well-organized union has actually simplified their labor problem. Certainly it is possible that a union might do a great deal to maintain the morale of workers, inculcate in them a sense of responsibility, improve workmanship, and reduce labor turnover.

It is commonly assumed that organized labor has substantially raised the general level of wages and living standards for workers, but whether this is really true is open to question. Standards of living have, to be sure, risen; but it is pretty clear that the explanation of this is to be found largely in improved methods of production, including the use of more and better machines. It is true that a strong union representing a limited group of workers can often secure substantial increases in both the money wages and the real wages of its own members. Because of this, most supporters of the labor movement assume that *all* workers, if they were only organized, could increase their real wages in the same way. It is not at all certain that this is so.

Since real wages consist of goods and services rather than money, there are only two ways in which organized labor could possibly raise the general level of real wages for all workers. The first is to cause a shift of income from the property-owning groups to the workers. Unions may be able to do this to some extent through collective bargaining, but for the economy as a whole the possibilities of shifting income in this way

are very limited. To begin with, the total income from property of people in the relatively high-income brackets is not sufficient to raise the average level of wages greatly even if all of it could be spread out among the workers. In the second place, most of this property income consists of rent and interest rather than profit in the strict sense, and collective bargaining is not an effective method of obtaining for workers a share of rent and interest. To the business firm, rent and interest are costs which in the long run must be covered if production is to continue. If unions should succeed in pushing wages high enough to encroach on normal payments of rent and interest, the effect would be to discourage production and decrease employment. The only other way in which employers could meet the situation would be to increase prices, and in that case the additional money received by workers in wages would be paid out to meet higher costs of living. As a result there would be no net gain.

The second way in which organized labor might conceivably raise the level of real wages for all workers is by bringing about an increase in the total output of goods and services available for consumption. It is quite possible that unions might be able to increase the efficiency and output of their members. It is doubtful, however, if they actually do this. In the majority of cases it seems more probable that they restrict output, in other words, that they seek more pay for less work. In this case, any real gain by a particular group of organized workers must be at the expense of someone else. Sometimes it may be at the expense of the employer. In the long run, however, much of it is likely to be at the expense of other groups of workers. These other groups bear the burden in one or both of two ways: they pay higher prices for goods produced by union labor, and in some cases they receive lower wages themselves. If they receive lower wages, it is because high union wage rates limit the demand for labor in the strongly organized occupations so that there is increased competition for jobs in those occupations in which labor is not organized.

Summary

After a long history of slow growth, the American labor movement achieved national importance when the Knights of Labor reached the height of its power in the 1880's. After the decline of the Knights, the American Federation of

Labor became our great national labor organization, but in recent years it has shared this position with the Congress of Industrial Organizations. The CIO came into existence as the result of a split in the AFL in 1935. The CIO was formed by a strong group of unions whose representatives believed that the mass-production industries, formerly unorganized, should be organized on industrial union principles.

Underlying the growth of modern labor unions was the industrial revolution, the development of the factory system, and the growth of separate capitalist and worker classes. Among the specific objectives that induce workers to organize are higher wages, shorter hours, better working conditions, greater security, and the desire to exert influence on legislation.

The basic unit of organized labor is the local union, but the center of power is usually the national union. A general federation is an association of different craft or industrial unions to further their common interests. Federations, which include such organizations as the AFL and CIO, have little power over their member unions.

When the representatives of a union and an employer negotiate, the process is called collective bargaining. If they call in a third party merely to give them aid and advice, they are said to resort to mediation; but if they decide to submit their dispute to a third party and abide by his decision, the process by which settlement is reached is called arbitration. The principal weapons of organized labor for enforcing its demands on employers are the strike and the boycott, reinforced by picketing. Organized labor is bitterly opposed to company unions which are dominated by the employer; and it generally favors the closed shop and the checkoff of dues because these tend to increase the membership and power of independent unions.

Antiunion activities of employers include the lockout, the blacklist, the yellow-dog contract, spying, use of strikebreakers, fomenting violence to discredit the union, and the court injunction. However, the ability of employers to use the injunction against workers has been greatly reduced by the Norris–La Guardia Act.

Under the New Deal, and especially as a result of the Wagner Act, organized labor grew in numbers and power. Another favorable piece of New Deal legislation was the Wages and Hours Act. During the war and postwar years labor continued to forge ahead, but during this period there was also a definite growth of antilabor sentiment among the public at large. This resulted in the passage by Congress of the Taft-Hartley Act, a statute that limits union activities in a variety of ways. It is too early to pass final judgment on this act. Some of its provisions may prove to have been unwise. Others, however, seem highly desirable from the standpoint of protecting the rights of both employers and the general public as well as those of the workers themselves.

The ways in which organized labor has influenced our economy are many. One effect has been the passage of much legislation favorable to workers. Another has been the creation of new types of problems for employers. Still another has been a substantial increase in the real wages of some groups of workers. However, the common assumption that organized labor has substantially raised the average level of real wages for all workers is probably false. Although increases in the general standard of living for workers have taken place, these increases seem to be attributable to other causes.

Study Questions

1. Why were the earlier successes of the labor movement confined to workers in the skilled trades?
2. What was one great weakness of the Knights of Labor?
3. What is the distinction between a craft union and an industrial union?
4. Explain the origin of the CIO.
5. Name several important unions not affiliated with either the AFL or the CIO.
6. Why does organized labor in the United States include only about 15 million workers out of a total labor force of some 60 million?
7. What are the principle reasons for the establishment and growth of labor unions?
8. How has free public school education affected the labor movement?
9. How are the following organizations related to one another: a local, a district council, a national union, a general federation, a city central? Which of these organizations is usually the center of power? Why?
10. Explain the terms collective bargaining, mediation, and arbitration.
11. Explain the difference between a primary boycott and a secondary boycott. Which is usually legal?
12. What is picketing, and when is it legal?
13. Why is organized labor strongly opposed to company unions?
14. Why do unions generally desire the closed shop and the checkoff of dues?
15. Explain what each of the following terms means: lockout, blacklist, yellow-dog contract, court injunction.

16. Why did organized labor regard the Norris–La Guardia Act as a legislative victory?
17. What were the purposes of the National Labor Relations Act? What were its principal provisions?
18. What are the principal provisions of the Fair Labor Standards Act?
19. Explain the rise of antilabor sentiment in the war and postwar periods.
20. What were some of the criticisms leveled at the Wagner Act?
21. What are the principal provisions of the Taft-Hartley Act?
22. Was organized labor justified in violently opposing the passage of the Taft-Hartley Act? Discuss.
23. What are some of the principal effects that organized labor has had on our economy?
24. Has organized labor raised the general level of real wages for all workers? Discuss.

References for Further Reading

Bye, R. T., and W. W. Hewett: "Applied Economics," F. S. Crofts & Co., New York, 1941, Chaps. 6 and 7.

Fairchild, F. R., E. S. Furniss, and N. S. Buck: "Economics," The Macmillan Company, New York, 1940, Chaps. 25-28.

Froman, L. A.: "Principles of Economics," Richard D. Irwin, Inc., Chicago, 1946, Chaps. 7 and 8.

Garver, F. B., and A. H. Hansen: "Principles of Economics," Ginn & Company, Boston, 1937, Chap. 28.

Gemmill, P. F., and R. H. Blodgett: "Current Economic Problems," Harper & Brothers, New York, 1939, Chap. 2.

Ise, John: "Economics," Harper & Brothers, New York, 1946, Chaps. 27 and 28.

Knight, B. W.: "Economic Principles in Practice," Rinehart & Company, New York, 1942, Chap. 5.

Mitchell, Broadus: "General Economics," Henry Holt and Company, Inc., New York, 1937, Chaps. 21 and 22.

27. Role of the Consumer

The consumer occupies an important place in the practical planning of production by businessmen as well as in economic theory. After all is said and done, we produce in order to consume. We may not consume exactly what we ourselves produce, but we engage in production in order to have something to exchange for the products that we require. To the man running a busines, the consumer is a tyrant whose every whim must be obeyed.

A word of warning to the student therefore seems in order. Care should be exercised not to minimize the importance of that phase of economics which deals with consumption merely because so little space is devoted to it in elementary texts. The economics of consumption is a complete subject in itself, but a subject that can better be understood after the student has obtained a thorough knowledge of economic principles. If the student continues his study of economics beyond the elementary level, he will have frequent contact with the economics of consumption.

The subject of consumption long has been one of the cornerstones of economics. As Alfred Marshall has pointed out, the older definitions of economics described it as a science concerned with the production, the distribution, the exchange, and the consumption of wealth. Under this definition, consumption would stand on an equal level of importance with the other divisions of economics that have occupied so much attention in this and other texts.

It is interesting to note the importance that Adam Smith attached to consumption. In his "Wealth of Nations" he states, "Consumption is the sole end and purpose of all production; and the interest of the producer ought to be attended to, only so far as it may be necessary for promoting that of the consumer." Kenneth Boulding, a contemporary American economist, has pointed out in the *American Economic Review* that "Under the stimulus of the emancipation of women and the de-

velopment of faculties of home economics, consumption economics almost threatens to become a separate science." In the writings of John Maynard Keynes and other modern economists, new emphasis on the economics of consumption has been given by the introduction of such useful terms as "propensity to consume" and the "marginal propensity to consume."

Further attention has been focused on consumption economics by the curious dilemma in which the world now finds itself. As a generality, the ability of consumers to acquire purchasing power has lagged behind our ability to produce goods and services. In other words, we have solved the problem of production better than the problem of getting the fruits of production into the hands of consumers. In order to provide employment for everyone, we must in the future make possible greater consumption than we have in the past. This certainly is true in the United States and probably will be true in many other parts of the world if peace prevails during the next decade. The real problem facing the United States and other nations to a lesser extent is how to place sufficient purchasing power in the hands of consumers so they can buy all the products that farms and factories can pour forth. We could, of course, approach the problem from another angle and ask, How can we make competition effective enough to reduce prices so that purchasing power will be adequate?

As important as the subject is, we cannot, however, devote much time to the economics of consumption. We shall even have to leave most of its basic propositions to more advanced courses. About all we can do here is to describe the role of the consumer and to point out the relationship between consumption and the other divisions of the science of economics.

The consumer and the price system. We already know that the desires of individuals appear to be insatiable in that all man's wants are never fully satisfied. We also know that satisfaction of these desires is curtailed by the limited purchasing power of individuals. Thus, the desire to consume is thwarted by inability of the individual to earn enough to purchase goods to satisfy all his desires. An individual always must compare the satisfaction that he expects to receive from consuming a good with the effort or sacrifice of leisure that he must put forth to earn the means of purchasing the good. The comparison between income on the

one hand and the price of a desired good on the other determines the quantity and variety of the goods that are consumed.

The fact that we live in an economy in which prices exercise control over production and distribution may well result in the bewilderment of the consumer. His confusion arises from the fact that he finds himself confronted with an infinite variety of goods available at constantly changing prices. These prices may increase or decrease in unison, but more commonly they fluctuate with reference to each other. For example, the price of wheat may increase while the price of electricity declines. Wages may remain constant while the prices of shoes, fuel, clothing, and other necessities change.

Not only is the individual consumer confused by price changes that affect his ability and willingness to buy goods and services, but also he seems helpless to do anything about it. It is easy for him to become convinced that sellers can set prices at will. When the buyer enters a store, he finds price tags on goods and he realizes that he either must pay the prices asked or leave without making a purchase.

The consumer, however, is not at all helpless in regard to prices. Although any one consumer cannot hope by his individual action to control prices, yet the combined action of consumers is of the greatest possible consequence in the price-making process. A producer quickly learns to his sorrow how easy it is to price his product out of the market. Waves of "buyers' strikes" have in the United States and in other countries brought down the prices of goods and services. We have seen that a monopolist or a producer operating in an imperfect market cannot control both the price of his product and the quantity that he sells. He may set the price, but consumers will then decide how much they will buy. On the other hand, he may resolve to sell a certain quantity, but in this case consumers will decide the price that they will pay. The rational businessman who has this choice will charge a price which permits the sale of the quantity that maximizes his profits. If he finds that he has set the wrong price, he will change it quickly.

The interesting thing about the role of the consumer in regard to prices is that the individual shopper usually does not cooperate consciously with his fellows in refusing to buy if prices seem too high to him. In periods of extremely high prices, such as have occurred after major wars, he may join an organized buyers' strike or a We Won't Buy It club. But for the most part he simply says, "No thank you!" when quoted

prices that he thinks are too high. Millions of other persons do the same thing, and soon prices come tumbling down. The individual need not feel chagrined at his apparent helplessness in the world of price tags, because in reality he is one of those who dictate prices.

Direct consumption and indirect consumption. In the preceding section our discussion of the relation of prices to the consumer was based on the assumption that goods were purchased for direct consumption. We were speaking of the reaction of the ultimate consumer to prices. However, many commodities are purchased because they are required in the process of production. The whole range of industrial raw materials, such as copper, pig iron, crude petroleum, and cotton fiber, falls into this classification. The using up of goods in the process of production is called *indirect consumption.* Industrial raw materials, for example, will be consumed ultimately by individuals in the form of finished products, but such consumption is indirect. The raw materials have to undergo further processing before they are fit for the final user. In summary, consumption is direct when a product is used by the ultimate consumer; consumption is indirect when a good is used up in making a product that will be subsequently used by the consumer.

We encountered closely related concepts in an earlier chapter when we distinguished between producers' goods and consumers' goods. Producers' goods needed further processing; consumption of producers' goods by the ultimate consumer was indirect. Consumers' goods were ready for immediate consumption; consumption of such goods was direct.

The consumption of some goods can be either direct or indirect, depending on how they are used. Coal is an example. As was pointed out, direct consumption of coal takes place when it is used for the heating of a home. Indirect consumption takes place when the coal is used to make steel, which in turn finds its way into an automobile. Businessmen are buyers of coal, as well as other goods consumed indirectly, and their demand is a vital factor in determining price. This means that the ultimate consumers are not the only persons who influence the prices of certain commodities that consumers use. A substantial portion of the demand for these commodities may come from businessmen.

However, even with respect to goods consumed indirectly, the role of the ultimate consumer is of major importance. Businessmen seek to acquire raw materials or semifinished goods as cheaply as possible in order to keep their cost of production low. They must keep an eye on

their marginal cost curves just as they study prices and marginal revenue, for their profits are maximized by equating marginal revenue and marginal cost. Thus they respond to the prices of the materials that they buy much as the ultimate consumer does. They refuse to buy if they consider prices are too high.

Foresight. Foresight is an important factor which affects both direct and indirect consumption. Those who purchase goods, either for direct consumption or for further processing, look to the future as well as the present. In the case of the ultimate consumer, limited purchasing power must be apportioned between unsatisfied present desires and prospective desires. Businessmen use goods in the present in the productive process in order that direct consumption may take place in the future. Therefore the businessman must incur costs in the present by purchasing raw materials, on the basis of his expectations of what consumers' desires will be when the finished product reaches the market. It is these future desires of consumers that will determine the prices that they will be willing to pay.

Foresight, therefore, is a very important consideration in business planning. A shrewd businessman gives much of his attention to anticipating the desires of consumers. Estimating the prices that consumers will pay in the future involves something more, however, than estimating their desires. It also involves what their incomes will be and what the prices of other goods will be.

Expectations of future prices are vital in business planning. Production requires time. Materials pour into the receiving department of a factory and sometime later the shipping department sends away the finished product. Still later, the consumer buys the product from a retailer. Months may elapse between the inflow of raw materials and the sale of the finished product.

A rational basis for planning exists if the businessman expects the price of the finished product to be the same as it is now. He can purchase raw materials now with considerable confidence. If he expects prices to be lower in the future than they are now, he has a serious problem in regard to planning his operations. He may hesitate to purchase raw materials because he cannot predict with assurance the extent of the decline that he believes will take place in the finished product. However, he may purchase raw materials now with confidence if he expects prices will rise.

Price expectations also concern the ultimate consumer. An individual will defer purchasing a product if he believes that it can be obtained at a lower price in the future. He will fill his cupboards and closets to the brim if he becomes convinced that prices will advance. The role of expectations in the sale of consumer goods can be seen from time to time when important changes in the price level appear probable. A good example was the experience of stores trying to sell women's fur coats in the winter of 1946-1947. Clearance sales of fur coats usually come in midwinter or later, but many stores reduced prices sharply even before Christmas. The price cuts, however, did not bring swarms of customers to the stores. The lack of enthusiasm of buyers at the time was attributed to expectations of still greater price cuts. During this same period, sales of many other items slowed in anticipation of price reductions. Similar reactions of consumers have occurred at other periods when a turn in the price level seemed probable. Any salesman knows from experience the role of price expectations of buyers.

Consumer choices. Although desires are unlimited, or at least incapable of complete satiation, man's productive power and nature's resources are not unlimited. Consumers, therefore, must choose between those goods which they feel they must have and those which they can do without. In other words, consumers have choices that should be made on a rational basis. The consumer should allocate his resurces, that is, his money, between possible expenditures in such a way that the satisfaction obtained from the expenditure of a dollar for any one commodity at the margin is equal to that obtained from the expenditure of a dollar for any other commodity. In other words, he should distribute his purchasing power between goods A, B, C, D, and E so that the marginal utility of a dollar's worth of any one of these goods is equal to the marginal utility of a dollar's worth of any of the others. This may involve purchasing more units of A than of B or C. It is the satisfaction obtained from the marginal dollar's worth of each good that he should equate.

Factors affecting the consumer's freedom of choice. A number of factors may modify the freedom that consumers have to choose the goods and services they purchase. Among these modifying influences are custom, the desire to display wealth conspicuously, lack of knowledge on the part of the consumer, fashion, marketing methods, and authoritative controls such as rationing. We shall consider each of these in turn.

Custom probably is the most significant single factor conditioning the desires of consumers. A custom has sometimes been described as a group habit. Consumers, for example, form the habit of purchasing certain kinds of goods. They expect to secure these goods at certain prices or within rather a narrow range of prices, and they may even confine their purchases to rather well-established quantities. The kind or type of food that a family buys often is determined in large part by custom. The preferences of certain segments of our population, established by custom, are given careful study by manufacturers and processors in an effort to cater to such tastes. Such preferences in regard to food may be local or national. Whatever the area, the fact that certain preferences are established by custom will affect vitally the choices of the consumer.

Custom also affects the demand for clothing. The dress of the average Chinese differs from that of the Arab or the American. A well-dressed person in one part of the world would be conspicuous in another. Notwithstanding changes in styles and fashions, evolution in clothing design is very slow, and present-day clothes include many features of the attire of former days. For example, the customary three-piece suit for men is said to be a carryover from an old military costume. A study of the history of clothing design reveals that clothes are affected by war, by religious practices, by fads and fashions, and only to a limited extent by the need of protection against the weather.

Custom affects the whole range of consumer choices. The way our homes are furnished often is dictated by custom. Even our choice of occupation may have been influenced by the fact that it has become customary in our family or group to enter the professions rather than the trades. We may be under pressure to continue in college because it has been the custom for members of our family to receive university training. In short, though consumer choices are not guided solely by custom, custom does have a powerful limiting effect upon the kinds of goods and services that consumers will purchase.

A second force affecting consumer choices has been called *conspicuous consumption,* or consuming goods in order to display wealth rather than to satisfy a real need. The practice of "keeping up with the Joneses" is rather well known. It is possible to do so by displaying financial well-being through conspicuous consumption or through the display of the leisure to engage in noneconomic activities. This consideration has often been pointed to as a factor in economic life. Thorstein Veblen, for ex-

ample, says in his book, "The Theory of the Leisure Class," "The basis on which good repute in any highly organized industrial community ultimately rests is pecuniary strength; and the means of showing pecuniary strength, and of so gaining or retaining a good name, are leisure and a conspicuous consumption of goods."

The desire to emulate the rich has many repercussions on the demand for various goods and services. It influences the demand for cheap furniture built to resemble its more substantial counterpart. The fur coat has become the symbol of wealth to the shop girl. The newly rich use the swimming pool, tennis court, and large home with many servants to impress their neighbors. Desire for display often asserts itself in the organization of exclusive clubs and in such things as travel, lavish entertainment, higher education, and philanthropy. Alert producers are often able to take advantage of this desire "to show off." They price goods used by the wealthy higher than otherwise would be possible because they know that a high price will make the good more desired. It has been contended that the well-advertised high price, rather than the odor, is responsible for the great demand by the feminine user, or her boy friend, for certain brands of perfume. Exclusive stores cater to the wealthy classes, and brands become the basis for determining whether a person belongs to "the right set." The fact that a hat was made by Bes Ben or that a woman's suit has an upper Fifth Avenue label is important in determining the satisfaction obtained therefrom, provided other women recognize that Bes Ben made the hat or that the suit really was purchased on upper Fifth Avenue. Conspicuous consumption, therefore, is a motive that cannot be ignored in an analysis of the pattern of consumer desires.

A third factor affecting the pattern of consumption is lack of knowledge on the part of the consumer. We have already pointed out that the consumer is confronted with an almost infinite variety of goods and is asked to choose among them. In many cases the consumer is in no position to make a rational choice. He does not have the time or ability, for example, to determine whether two suits priced approximately the same are equal in quality. He cannot always tell whether a garment advertised as being all silk actually contains only silk. The size of the package that he purchases may be deceptive or the contents of packages that appear to be the same size may vary.

The inability of consumers to choose wisely because of such considerations enables producers to deceive customers if the producers are unscrupulous, even though the quantity of the contents and the nature of the ingredients may be stated on the package. Under such circumstances, the consumer tends to depend on price as the criterion of quality. The usual economic process is in a sense reversed when consumers lack knowledge. Instead of quality controlling price, price becomes the basis for judging quality; in other words, price determines utility in the mind of the purchaser.

A fourth factor affecting consumption is the power of style or fashion. Much of the power of fashion in influencing purchases arises from the desire to imitate those who are presumed to be the leaders in the community. Therefore the desire to be in fashion is somewhat related to the desire for conspicuous consumption. The power of fashion over consumers also grows out of the perpetual desire for something new and different.

In some respects custom and fashion are contradictory influences. The essence of fashion appears to be change, while custom resists change. As a matter of fact, custom does prevent too radical a change in fashion in short periods of time. In many instances a change in fashion may not be to something new; many times it represents going back to a fashion that prevailed at an earlier time. The change in the length of dresses is an example of how style changes travel in cycles.

It is clear that a change of style in clothing, even if slight, serves as a powerful stimulant to sales, since old wardrobes become obsolete. The same principle applies to changes in automobile models. Hence manufacturers often encourage such changes to create new demand for their products. However, well-established consumer customs and tastes set limits to such changes.

A fifth factor affecting consumer choices is marketing methods. This term includes all selling efforts of producers or their sales representatives to influence the choices of consumers. We must include under marketing methods the kinds of outlets through which goods are made available to the public, and the terms of payment offered. Of the various types of selling efforts, we shall first consider advertising.

The amount spent for advertising is enormous. In the United States it runs from about 2 to 3.5 billion dollars a year, depending on the degree of prosperity that the nation is enjoying. A businessman may have one

or more objectives in mind when he advertises. He may be seeking to inform consumers of the existence or uses of a new product in order to win customers for it. He may, however, merely be seeking to induce consumers to use his brand instead of some other. Or again, his advertising may be almost entirely defensive in the sense that he is trying to prevent other producers from taking his customers.

Some advertising is educational. Many businessmen are convinced that consumers do not know definitely what they want or what products are available. By giving the public information on new and better products, the businessman benefits consumers and at the same time increases the sales of his product. The same principle applies when the businessman uses advertising to point out new uses for old products.

Advertising used to inform the public about new products and to point out new uses for old products is socially useful. It is conceivable also that advertising may benefit the consumer by making it possible to give him products at lower cost despite the expense of the advertising. This would be the result if the advertising so increased the demand for a product that the manufacturer was able to produce at lower unit cost. For this benefit to accrue to the consumer, the businessman would have to pass along some of the savings in cost in the form of a lower price for the product.

There appears to be somewhat less social justification for advertising designed merely to shift consumers' purchases from one company to another rather than to create new markets. Advertisers of cigarettes spend millions of dollars every year for radio programs and newspaper advertising primarily to convince smokers of the advantages of their particular brands.

Critics of advertising contend that such advertising is designed to limit freedom of choice by consumers rather than to better qualify consumers to make choices. In other words, it is charged that advertising through persuasion helps to make up the consumer's mind for him in the way that the advertiser wants it made up.

There are other phases of marketing procedure which involve the methods used by producers to attract customers. One method of winning customers is attractive and convenient packaging. Another is the making of buying easier by the extension of credit. Each of these methods warrants further discussion.

The art of packaging is an important one in selling goods. Packaging goods in attractive containers and in convenient quantities may increase prices, but the consumer is willing to pay extra because he receives additional satisfaction. Demand may be shifted from one producer to another simply because one producer presents his goods in a better manner than his competitors. The alert manufacturer of consumers' goods will invest substantial sums in research to determine the kind and size of package that the customer prefers.

The use of credit probably is even more significant as a marketing device. Businessmen extend credit to consumers for either or both of two reasons. In the first place, the buyer may be unable to pay the entire purchase price of the goods at the time of sale. This often is true when larger items are purchased, such as furniture, automobiles, and houses. A second reason for extending credit is to add to the convenience of the customer. When credit is available, the consumer does not have to carry cash or write a stream of checks. He waits until the end of the month and then pays for all purchases made during the period. This method of paying for his purchases aids the consumer in keeping a record of his expenditures.

Installment selling, which is a special method of credit extension, has exercised an important influence on the demand for durable consumers' goods. Although installment selling has been used for almost 150 years in the United States, its real significance was not appreciated until it began to be used extensively in the selling of automobiles in the early 1920's. The special object of installment selling in this field was to tap the market represented by the lower income groups. The success of the automobile industry in selling cars to persons of low incomes soon led to installment selling in such fields as radios, furniture, washing machines, refrigerators, and clothing. As a result, the volume of such sales has increased tremendously. In some prewar years, 10 per cent or more of total sales of goods at retail were made on a monthly payment basis.

The popularity of the installment plan is not hard to explain. Because of preference for the present over the future, some persons find saving difficult. The installment plan permits such individuals to have expensive goods now without going through the painful process of saving first. They claim that it is less painful to pay while you enjoy the goods. All they must have at first is the down payment. A new selling weapon thus is placed in the hands of the businessman. He can point out how easy it

has become to buy. You merely put up a small down payment, and you have the good.

The activities of small loan companies and credit unions are also a factor in temporarily increasing the purchasing power of consumers. The ability to borrow substantial sums and to repay on the installment plan helps the members of low-income groups to meet emergencies. Low-income groups have inconsequential reserves for coping with such emergencies as sickness, accident, or unemployment. About their only source of assistance at such times is a lending agency.

Installment selling can be a force for good when it enables people to buy urgently needed goods and services in the present instead of postponing their acquisition to the future. However, it cannot be denied that the present increase in purchasing power is usually at the expense of a decrease in future purchasing power. Often installment selling has been abused. In some cases the cost to the consumer of buying on the installment plan has been pushed up far higher than would be warranted by reasonable rates of interest or the actual costs involved in extending credit. The justifiable charge also is made that installment selling is a method of high-pressure selling when in the hands of unscrupulous merchants. Fortunately for consumers and for society, the better merchants are too intelligent to oversell the customer. They know that they may have to take the goods back if the customer buys more than he can pay for in monthly installments. At best, however, the consumer pays more when he buys on the installment plan, and this tends to reduce his ultimate total purchasing power.

Authoritative control. Government regulation of the production and sale of goods may effectively limit the freedom of consumers to choose the goods and services they want. During the Second World War the peoples of nearly every nation had to accept rationing of many commodities in the interest of national defense. Rationing was much more severe in some countries than in others. Some countries had to extend rationing far into the postwar period because of the dearth of raw materials and the need for repairing factories and transportation systems destroyed by bombing.

The question of government control of consumption in normal times is quite another matter. In the period before the Second World War we saw many examples of the control by authoritarian government of the goods and services desired by their citizens. There are many ways in

which such control can be exercised. An authoritarian government can restrict imports. By decree it can forbid the production of certain commodities. In an extreme case it can even acquire the means of production.

When such controls are exercised by government, the question is not, What does the consumer want? Rather, the question becomes, What should the consumer have? We shall not take time here for detailed comparison of the position of the consumer under communism, socialism, fascism, and our own system of free enterprise. However, we can point out that our free-enterprise system regards the consumer as king; the other systems look upon him as a very unimportant fellow.

In the United States, under our free-enterprise economy, the efforts of three groups are directed at improving the position of the consumer: (1) the attempts on the part of the government to assist the consumer, (2) action on the part of consumers working in cooperation to protect their own interests, (3) action on the part of producers to protect consumers' interests.

In the main, the government's concern for the consumer has been to protect him from the practices of certain producers. The government seeks to secure a fair price for the consumer by preventing monopoly and combinations in restraint of trade. The consumer is also protected by laws against price discrimination. Misbranding, adulteration, and misleading advertising are forbidden under the Pure Food and Drugs Act, the Clayton Act, the Wheeler-Lea Act, the Wool Products Labeling Act, and other legislation. These laws forbid certain acts of producers in order to protect consumers; they do not really restrict the choices of the consumer. He still has freedom of choice, but his selection is limited to goods which are pure or which are clearly labeled as to weight, content, and ingredients.

Only when the government absolutely forbids the use of certain goods is the choice of the consumer limited. When the prohibition amendment was passed, the consumer was forbidden to purchase liquor and the producer was forbidden to sell it to him. The sale of opium has long been forbidden. Fortunately the list of things that the consumer is forbidden to buy is rather limited, and most restrictions have been placed, not on the consumer per se, but rather upon the producer in order that the consumer may be protected. As yet authoritative control has not much limited the freedom of choice of consumers in the United States.

The government, however, has not been satisfied simply to obtain

passage by Congress of a few laws regulating production in the interest of the consumer. A bureau of standards has been created to maintain standards of weights and measures. Other standards have been established and applied to many items sold in interstate commerce.

Furthermore, various departments of government have established bureaus to aid consumers. The Department of Agriculture has several of these bureaus, such as the Bureau of Home Economics, the Bureau of Agricultural Economics, and the Office of Information. These offices gather information and release it to the public so that the consumer may be better informed. Other departments also contain agencies whose purpose it is to keep the consumer informed.

State governments have carried on such work, as have also larger cities. Thus, the consumer has really been aided by government in the United States rather than restricted.

In addition to the aid given the consumer by the government, other agencies also assist him. Most newspapers will carry no advertising that they suspect is fraudulent. Many patent medicines have been refused space for this reason. The Better Business Bureaus found in most cities are organizations that seek to prevent illegitimate selling methods or the marketing of misrepresented products. The Good Housekeeping Institute, formed in 1912, has become the guide for many consumers. Trade associations often establish testing laboratories where goods are compared with rigid quality standards.

Some merchants protect consumers by testing the goods that they purchase. The consumer, knowing that the merchant is purchasing only goods of high quality, can buy such items with confidence.

There is still another important form of aid to the consumer. This is afforded by the producer himself. By use of brands and trade-marks, producers clearly identify their products. A consumer who buys trade-marked or branded goods can subject them to rigid tests. If such goods prove to be as represented, the consumer can continue to buy; if not, the producer has lost a customer. Since the producer stakes his economic future on his product when he identifies it by label or trade-mark, he is likely to try to keep the quality high.

The choice between spending and saving. In addition to selecting the goods and services that he wishes to purchase, another choice of greatest importance is open to consumers, particularly those in the upper income groups. The consumer decides what proportion of his income he will

save and what proportion he will spend. Such a choice is very much restricted for low-income groups. In the very lowest income groups, all income must be spent and the government or private agencies have to subsidize additional consumption. But on an over-all basis, consumers have considerable latitude in deciding what shall be spent and what saved. Table 32 shows how Americans have allocated their incomes between saving and spending in some recent years.

Table 32.—Disposable Income, Consumer Expenditures, and Net Savings of Individuals, 1939–1946 (Billions of Dollars) *

	1939	1940	1941	1942	1943	1944	1945	1946
Disposable personal income	$70.2	175.7	$92.0	$116.2	$131.6	$146.0	$150.7	$158 4
Personal consumption expenditures	67.5	72.1	82.3	90.8	101.6	110.4	121.7	143.7
Personal savings	2.7	3.6	9.8	25.4	30.0	35.6	29.0	14.8

* Source: Department of Commerce. Rounding of first decimal accounts for failure of personal savings to equal precisely the difference between disposable income and consumption expenditures in all years.

During the war years, the scarcities of goods and the war bond sales restricted somewhat the freedom of choice between spending and saving. However, the scarcities of goods were more important in changing the pattern of spending than in curtailing the freedom to decide whether to spend or save.

Propensity to consume. It makes little difference to a wealthy nation such as ours if a particular individual decides that henceforth he will save more and spend less. However, such a decision on the part of great numbers of individuals could create havoc. Merchants would become panic-stricken as customers stayed away; manufacturers would receive few new orders and many cancellations of orders already placed would pour in, because merchants would seek to work off inventories. Factories would close and unemployment would skyrocket. Incomes would fall so rapidly that the individuals who planned to increase their savings might be frustrated. Because of lower incomes, they could not save as much as they had intended.

The dire sequence of events just described might not occur if the increased savings of individuals could find immediate investment in industries producing capital goods. The construction of capital goods would provide employment to offset the fall in employment offered by makers of consumers' goods. Because income earned in producing capital goods would be spent for consumers' goods, the consumers' goods industries would continue to be active.

During the depressed 1930's, and at many other periods, there was not an immediate outlet for savings in capital goods industries on terms that lenders were willing to accept. Hence, the decisions of individuals as to what proportion of their income they would spend and what proportion they would save became very important. That proportion of an individual's income which he elects to spend is a measurement of what economists sometimes call his *propensity to consume.* During periods of restricted investment opportunity, a high propensity to consume on the part of all individuals means more employment in consumers' goods industries and hence a higher level of over-all employment.

The lower income groups automatically have a high propensity to consume. As the wealth of an individual increases, his propensity to consume declines. A decreasing proportion of his income must go to acquire the necessities of life. As the wealth of a whole nation increases (and wealth and income have risen in most western countries), the falling propensity to consume may make attainment of full employment increasingly difficult except during periods when abundant and acceptable investment opportunities are at hand. This is especially likely to be true if competition becomes more imperfect and the economy becomes more rigid.

This brings us to another very useful term, "marginal propensity to consume." The proportion of any increment of income that will be consumed measures the marginal propensity to consume. A poor man will consume all or almost all of any increase in income. A wealthy man may consume almost none of his increment in income. In trying to appraise what portion of income will be saved and what will be spent as the wealth of a nation increases, the significant consideration thus is the marginal propensity to consume.

Earlier in this chapter the comment was made that the United States, and possibly some other nations, have solved the problem of production better than the problem of maintaining high purchasing power. The

concepts of propensity to consume and marginal propensity to consume throw some light on the reasons for this. When a nation has become very productive and very wealthy, large incomes are generated through the process of production. A low marginal propensity to consume, however, means that a relatively small portion of an increase in incomes will be used for consumption. Hence, only part of large incomes will be used up in consumption. The remainder will be saved instead of spent for goods and services, the production of which provides employment. As we have already pointed out, the result is a tendency toward falling incomes and rising unemployment unless price adjustments take place quickly, so that consumption is stimulated and new outlets for the savings are found in capital-producing industries. As already indicated, this has not always been possible in the past because from time to time we appeared to have constructed all the capital goods that could be employed profitably. Thus it becomes evident that we are faced in the United States with the necessity of consuming more than we have in the past. This of course means only that we are faced with the necessity of raising our real standards of living. Such a situation is in itself highly fortunate. The problem, however, is so to organize our economy, so to adjust prices and incomes, as to make it possible for an ever-increasing output of goods to flow smoothly into the channels of consumption.

Summary

This chapter has pointed out factors that affect the role of the consumer in our economic system. We know that production is not an end in itself and that all production has for its ultimate purpose the satisfaction of the desires of consumers. However, we have learned that the free choice of consumers is restricted in many ways: by custom and habit, by lack of knowledge of conditions in the market, by the desire for display, by fashion, by marketing methods, and by state intervention. Various attempts to aid consumers were discussed. We noted that often consumers cooperate to aid one another; that producers sometimes help consumers to purchase wisely; and that the states and the Federal government provide information and certain types of protection for consumers. In spite of these efforts to aid the consumer, he is as yet a long way from attaining his goal of being able to purchase goods intelligently. We also pointed out in this chapter the economic repercussions that may result from decisions of consumers to change the relative proportions of their incomes which they devote to spending and saving.

Study Questions

1. In what way does the consumer exert a controlling influence upon the economy?
2. "It is clear that the individual consumer cannot influence prices; as a result, prices usually are determined by producers." Comment.
3. What is meant by saying that consumer influence in prices is indirect?
4. What is the difference between direct consumption and indirect consumption?
5. Name the factors that affect the consumer's freedom of choice.
6. What are the advantages claimed for installment selling? What are the possible abuses?
7. What methods are advocated to aid the consumer?
8. If you were asked to devise ways and means to aid consumers, what suggestions would you make?
9. What is meant by the propensity to consume? Explain the economic effects that may result from a change in the propensity to consume.

Exercises

1. Study the advertisements found in any magazine. Determine carefully the type of appeal that each presents. List these and point out whether the appeals are beneficial to the consumer.

2. Study some of the reports published by the Federal Trade Commission and point out the methods or devices used to secure consumer approval of the manufacturer's product.

References for Further Reading

Bye, R. T.: "Principles of Economics," F. S. Crofts & Co., New York, 1941, Chap. 24.

Gemmill, P. F.: "Fundamentals of Economics," Harper & Brothers, New York, 1943, Chap. 30.

KYRK, HAZEL: "A Theory of Consumption," Houghton Mifflin Company, Boston, 1923, Chaps. 2-5.

LEVEN, MAURICE, and OTHERS: "America's Capacity to Consume," Brookings Institution, Washington, D. C., 1934.

MARSHALL, A.: "Principles of Economics," Macmillan & Co., Ltd., London, 1930, Book VI, Chap. 13.

MCCONNELL, D. W., E. AYRES, A. A. FRIEDRICH, and W. E. ATHENS: "Economic Behavior," Houghton Mifflin Company, Boston, 1939, Chaps. 34 and 35.

MOORE, J. H., and M. L. HOCK: "Readings in Modern Economics," Thomas Nelson & Sons, New York, 1941, Chap. 8.

REID, MARGARET G.: "Consumers and the Market," F. S. Crofts & Co., New York, 1938.

WYAND, C. S.: "The Economics of Consumption," The Macmillan Company, New York, 1937.

Index

A

Ability-to-pay principle, 399
Accountant profit, 386
Advertising, 440
Allied Chemical and Dye Corp., 246
American Economic Review, 432
American Federation of Labor (AFL), 416
  and CIO, 417, 419
  organization of, 415-417
American Smelting and Refining Co., 251
American Telephone and Telegraph Co., 43
American Viscose Corp., 246
Average cost, 158
  in intermediate period, 156, 166-172
  in long run, 193
  lowest, 176
  relation to fixed cost, 158
  relation to variable cost, 158
Average product, 348-349
Average revenue, 223

B

Balanced consumption, 14
Balanced economy, 296
Bank, commercial, 64
  Federal Reserve, 65
  history of, 65
  member, 73
Bank for Reconstruction and Development, 282
Bank check, 57
Bank deposits, creation of, 69
  process of, 69-72
Bank money, 55
Bank notes, 55
Bank reserves, 64
  required, 67
Banking systems, 65-66
Barter, definition of, 50
  obstacles to, 50
Benefit principle, 399
Best combination principle (principle of proportionality), 172
Bethlehem Steel Corp., 250
Better Business Bureaus, 445
Bills of exchange, 288
  buyers and sellers of, 291
Bimetallism, 52
  history of, 53
  and market ratio, 53
  and mint ratio, 53
Board of Governors, Federal Reserve, 66
Bonds, corporation, 45
  government, 66
Boulding, Kenneth, 432
Brotherhood of Locomotive Engineers, 419
Bulk-line cost, 215-217
Bulk-line price, 217
  theory of, 217
Bureau of Labor Statistics, 89-91
Business activity, 16
  based on self-interest, 16
  essential conditions for, 6-7
  forms of, 43
  government control of, 17
  large scale, advantages of, 40
  limitations of, 41
  organization of, 42

Business activity, problems of, 8
social implications of, 2
tests of, 4
Business cycles, 106, 108
Business enterprises, size of, 39
number of, 43
Businessman, 1
relation to society, 15

C

Capital, 30, 365
demand for, 368-371, 375
infinitesimal units of, 330*n*.
and interest rate, 376
liquid or free, 365
supply of, 371
time preference, relation to savings, 372
Capitalism, 18
Capitalistic system, 18
Celanese Corporation of America, 246
Central banks, 66
Chandler, Lester V., 85
Chrysler Corp., 152
Clark, J. M., and government regulation of economic activity, 16
eleven rules for, 18
Class price, 233-239
Clayton Act, 244
Closed shop, 421
Coinage, free, 52
Coins, token, 58
Colbert, Jean Baptiste, 277
Collective bargaining, 420
Commercial banking, 64
Commercial paper, 66, 68, 72, 73
Committee for Industrial Organization, 416
Common stock, 45
Company union, 421
Comparative advantage, principle of, 283-287
Competition, advantages of, 20
and character of economy, 18
Competition, conditions necessary for, 139
disadvantages of, 20
and duopoly, 245, 248
effect on employment, 106
on goods, 143
imperfect, 144, 244-256
monopolistic, 244, 254-256
and oligopoly, 246-248, 250-254
perfect, 138-140
pure, 139*n*.
Concepts, fundamental, definitions of, 28-32
Congress of Industrial Organizations (CIO), 416-419
Conspicuous consumption, 438
Constant-cost industries, 198
Consumer, 23, 432-449
advertising for, 440
aid to, 445
choice between spending and saving, 445
and custom, 438
and installment selling, 442
marketing methods for, 440
packaging, effect on, 442
and price system, 443
and style, 440
Consumer choices, 437-446
Consumer credit, 442
Consumer goods, 29
Consumption, balanced, 14
conspicuous, 438
direct, 435
indirect, 435
propensity to consume, 102, 446-448
related to savings and investment, 100
Contraction, cyclical, 107
Cooperation, voluntary, 18
Corning Glass Works, 245
Corporation, advantages of, 45
definition, 45
disadvantages of, 46
Corporation bonds, 45
Cost, average, 156, 158, 166-172, 176, 193

Cost, differences in, 187
and diminishing returns, 170
fixed, 157
future, 156
in intermediate period, 166-172
long-run, 193-194
marginal, 157
opportunity, 155
past, 156
and price relationships, 155
short-run, 155
variable, 157
Cost curve, of average firm, 166
Credit expansion, 69
Credit money, 69
Currency, 30, 50
(*See also* Money)
Customs duties, 408
Cyclical contraction, 107

D

Decreasing-cost industries, 199
Definitions of fundamental concepts, 28-32
Degressive taxation, 402
Demand, changes in, 119
definition of, 115
and desire, 115
and diminishing utility, 116
elasticity of, 121
equilibrium with supply, 135
for labor, 307, 309
nature of, 114
for product of individual producer, 185
and satisfaction, 116
static concept of, 309*n*.
Demand curve, 119
Demand schedule, 115-118, 183
changes in, 137
individual, 117
market, 117
Deposits, bank, 57, 69
derivative, 78
primary, 78
Desires, economic, 13
noneconomic, 13
Determination, simultaneous, 313
Diminishing returns, average, 169
definition of, 168
law of, 15
marginal, 170
principle of, 167, 286
relation to cost of production, 170
Diminishing utility, principle of, 13, 116
Direct consumption, 435
Disposable income, 436
Distribution, 302-316
as branch of price theory, 303
and determination of shares, 310-312, 313
meaning of, 302
relation to productivity, 305
marginal, 306-308
Distribution theory, 313
District council, 419
Dollar, U.S., 52-55, 57-60, 273, 277-279, 289
Dun and Bradstreet, Inc., 388
Duopoly, 245, 248-250
Du Pont de Nemours, E. I. & Co., 246
Durable goods, 133
Dynamic economy, 206-217
business adjustments under, 211
definition of, 206
equilibrium prices in, 209-211
long-run equilibrium in, 207
long-run price adjustments in, 206
price determination in, 207-211
predictions in, 211

E

Economic goods, nonmaterial, 29
Economic harmonies, 15
Economic instability, 108
Economic limits to taxation, 403
Economic principles (*see* specific principle)
Economic problems, 9

Economic society, 18
Economics, definition of, 1
  foundations of, 13
  purpose of study, 9
Elastic money, 68
Elasticity of demand, 121-124
Employment, area of, 107
  effect on income and expenditures, 103
  effect of speculation on, 105
  full, basic conditions of, 109
  increases in, affected by competition, 103
  level of, 106
Enterprise, 39
  free, 18
    (*See also* Business Activity; Business Enterprise)
Entrepreneur, definition of, 1
  history of, 38
  identification of, 382
  role of, in free society, 382
  services of, 39
    payment for, 302
  as stockholder, 390-392
Equal-sacrifice principle, 400
Equation of exchange, definition of, 82
  formula for, 82
  limitations of, 82
Equilibrium of demand and supply, 135-144
Equilibrium adjustments, 207-211
Equilibrium price, and changes in schedules, 137
  determination of, 136-137
  during intermediate period, 141
  during long run or normal, 141
  modifications of, 139
  during short run, 140
Exchange, indirect, 291
Excise taxes, 406
Expenditures, relation to income, 99
  national, items included in, 103*n*.
External economies, 199

F

Factors of production, 30, 308-314
  contractual price of, 312
  demand for, 308, 322
  prices of, 304, 310
    determination of, simultaneous, 313
  supply of, 309
  unlike units as, 312
Fair Labor Standards Act of 1938 (Wages and Hours Law), 423
Federal Reserve Act of 1913, 55
Federal Reserve agent, 68
Federal Reserve Bank, 68
  reserve ratio, relation to legal ratio, 67
Federal Reserve bank notes, 56*n*.
Federal Reserve bank reserves, 67
Federal Reserve Board of Governors, 66
Federal Reserve Bulletin, 58
Federal Reserve notes, 59, 68, 73
  security for, 68
    commercial paper, definition of, 68
Federal Reserve System, 66
Firestone Tire and Rubber Co., 253
Firms, high-cost, 213
Fixed costs, 157
Fixed stock, 222
Ford, Henry, 124
Ford, Henry, Jr., 177
Foreign trade (*see* International trade)
Foresight, 13, 372
Free coinage, 52
Free enterprise, 18
Free trade, 276
Future costs, 156

G

General Electric Co., 245
General federation, 419
"General Theory," 100*n*.
General Tire and Rubber Co., 253
General welfare, 24
Glass-Steagall Act of 1932, 69
Gold certificates, 58*n*., 68

Gold money, 54
Gold points, 288-291
Gold Reserve Act of 1934, 57
Gold standard, 51, 290, 294
  advantages and disadvantages of, 54
  and stability of exchange rates, 54*n*.
Good Housekeeping Institute, 445
Goodrich, B. F., Co., 253
Goods, capital, 310, 321
  consumers', 29
  definition of, 28
  economic, 28
  free, 28
  producers', 29
Goodyear Rubber Co., 253
Government, 17
  costs of, 395-397
    apportioning of, 399-401
Government bonds, 66
Government control of consumption, 443
Government debt, 410
Government money, 55
Government revenue, 397, 408
  tariffs, 406
  taxes, 406-408
    income, 405
Grants from Federal government, 408
Greenbacks, 54
Gresham's law, 53
Gross national product, definition of, 396*n*.

H

High-cost firms, 213
Hoarding, 101, 101*n*.
Home market, 296
Hume, David, 85

I

Imperfect competition, definition of, 244-256
Incentives, effect of taxes on, 403
Income, definition of, 30
  disposable, 436
Income, national, 314-316
  psychic, 346
  relation to cost, 155
  relation to expenditure, 99
  relation to speculation, 105
  taxes on, 316, 405
  types of, 30
Increasing-cost industries, 197
Independent union, 421
Index numbers, definition of, 89
  and weighted aggregates, 89
Indirect consumption, 435
Indirect exchange, 291
Individual producer, 185-188
  cost curve of, 185-186
  demand for product of, 185
  effect of price changes on, 187
  in the intermediate period, 185-188
  marginal cost of, 188
  output of, 186
Industries, constant-cost, 198
  decreasing-cost, 199
  increasing-cost, 196
  infant, 295
  profits in U.S., 387-389
Inelastic demand, 121
Inelastic money, 65
Inseparable-cost goods, 267
Installment selling, 442
Interest, 365-377
  contractual, 367
  definition of, 366
  determination of, 373
  pure, 367
  rate of, 367, 373-377
Intermediate period, 141
  definition of, 183
  demand schedules in, 183
  differences in average costs in, 187
  effect of price changes, 187
  equilibrium price in, 186
  individual firm in, 185-188
  marginal costs and price in, 186, 188
  monopoly price in, 225

Intermediate period, price in, 183-191
relation to cost, 186
supply in, 184
Internal economies, 199
International Bank for Reconstruction and Development, 280, 282
International Harvester Co., 251
International Ladies Garment Workers, 419
International Monetary Fund, quotas of, 281, 295
International trade, 273-298
advantages of, 276
agencies fostering, 280-282
attitude toward, 276
basis for, 282
domestic versus international, 273, 275
mechanics of, 287-290
and principle of comparative advantage, 283-287
stake of U.S. in, 277-279
Investment, relation to saving and consumption, 100
Invisible hand, 16
Invisible items, 292

J

Joint-cost products, 265-267
Joint demand, 261-263

K

Keller, K. T., 152, 162
Keynes, J. M., theories, on equality of investment and saving, 100*n*.
"General Theory," 100*n*.
on liquidity preference, 105
on propensity to consume, 102, 433
on unemployment, 108, 108*n*.
Knights of Labor, 416

L

Labor, 30, 313, 315
and competition, 350
Labor, demand for, 347
as factor of production, 30, 341
immobility of, 343
infinitesimal units of, 330*n*.
and law of substitution, 350
marginal productivity of, 307, 348
under New Deal, 423
and noncompeting groups, 344
supply of, 348-350
and unemployment, 314
war and postwar, 424
Labor market, 341-343
and limitations on competition, 350
Labor monopoly, 357
Labor organization, 415-428
development of, 415-417
district council in, 419
economic effects of, 427
general federation in, 419
growth of, 417-419
structure of, 419
weapons of, 420
Labor unions, 313
and antiunion activities, 422
closed shop versus open shop, 421
and collective bargaining, 420
company versus independent, 421
craft, 416
growth of, 417-419
industrial, 416
local, 419
national, 419
wage theories of, 357
Laissez faire, 17
Land, 30, 320
capital value of, 334
demand for, 324
extensive-intensive cultivation of, 326, 327, 330
formula for rent, 332
marginal, 329
marginal productivity of, 326
submarginal, 329
supply of, 324

Land, unlike units of, rent on, 326
Large-scale enterprise, advantages of, 40
  definition of, 39
  limitations of, 41
Law, of diminishing returns, 15, 306
  of one price, 139
  of substitution, 350
Legal limits to taxation, 403
Legal reserves, 64
Legal tender, 67
  (*See also* Money)
Lewis, John L., 416, 417, 424
Limited liability, 45
Limits to taxation, 402-403
Liquidity preference, 105
Long-run period, 141
  average-cost curve for, 194
  cost trends for industries, 196-200
  cost trends in, relation to price, 200-202
  nature of, 192
  price in, 192-203
  and size of plant, 193, 194, 196
Lowest average cost, 176

M

Malthus, Thomas, 353
Malthusian theory, 353
Man, nature of, 12
Management, 37
  wages of, 384
Marginal cost, 157, 174
Marginal price offer, 117
Marginal product, 169, 349
Marginal productivity, 306
  of labor, 307
Marginal productivity schedule, 306-308
Marginal propensity to consume, 447
Marginal revenue, 174, 220, 224, 228, 232
Market demand, for factor of production, 308
Market demand schedule, 117
Market demand schedule, changes in, 120
Market price, 141
Market ratio, 53
Market sharing, 252
Market supply, of factor of production, 309
Market supply schedule, 132
Marshall, Secretary of State, 278
Member banks, 73
  (*See also* Banks)
Mercantilism, 277
Monetary standards, 51-54
Monetary unit, standard, 52
Money, absolute value of, 81
  bank, 55
  bank deposits as, 57
  changes in value of, 81
  cheap, 53
  definition of, 30, 50
  determination of value of, 81
  effect of changes in value, 91
  elastic, 68
  functions of, 51
  government, 55
  as legal tender, 67
  newly created, 60
  paper, 54
  pocket, 59
  purchasing power of, 91
  quantity theory of, 85
  standard, 51
  till, 66
  in use, 58
  velocity of, 82, 84, 99, 109
Monopolistic competition, 244, 254-256
Monopoly, 19-24
  advantages of, 22
  characteristics of, 18
  class price, 233-239
  partial, 244-254
Monopoly price, under constant cost, 228
  under decreasing cost, 231-233
  determination of, 221
  under increasing cost, 230

Monopoly price, in intermediate period, 225
long-run, 227-233
short-run, 224
Multiplier theory, 104, 109

N

National Banking Act of 1863, 55
National economy, instability of, 108
National income, concept of, 99*n.*
contraction and expansion in, 155
division of, 314-316
formula for stability of, 99
proportion spent on consumption, 100*n.*
National Industrial Conference Board, 106
National Labor Relations Act (Wagner Act), 423
National Labor Relations Board, 423, 426
New Deal, 417, 423
Noncompeting groups, 344-347
Nonmaterial economic goods, 29
Normal price, 141*n.*
Norris-La Guardia Act, 421, 423

O

Oligopoly, 246-248, 250-254
Open shop, 421
Opportunity costs, 154
Organization, 40-47
Organized society, 15
Output, relation to fixed and variable cost, 158

P

Paper, commercial, 68
Paper money, 53
Paper standard, 54, 58
Partial monopoly, 244-254
Partnership, 44
Perfect competition, 138-140
equalization of price and marginal cost under, 186, 188
(*See also* Competition)
Petrillo, James V., 424
Pocket money, in circulation, 59
Political limits to taxation, 403
Preferred stock, 45
Price, and cost of production, 150
under dynamic conditions, 206-218
equilibrium, 135
under joint demand, 261-263
law of one price, 139
normal, 141*n.*
and production, 150
variation in, 103
Price determination, intermediate period, 141
long-run, 141
relation to supply, 131
short-run, 140
Price discrimination, 234
Price level, measuring changes in, 88
by index numbers, 89
by weighted aggregates, 89-90
Price offer, marginal, 117
Price system, 260
Prices, of inseparable-cost goods, 267
interdependent, 260-289
of joint-cost products, 265-267
of substitutes, 263
of unlike factors, 312
under joint demand, 261-263
Primary deposits, 78
Principles, economic (*see* name of specific principle)
Producer, 28
Product differentiation, 254-256
Production, in business, 40
cost of, 150
definition of, 29
and diminishing returns, 170
factors of, 30, 308-314, 322
nature of, 35
variations in, 103
Productivity, 305-308
of labor, 351
marginal, 306-308

Productivity, as source of demand for factor of production, 305
Profits, accountant's views on, 386
in American industry, 387-389
average, 158
corporate, 315
criticism of, 389
definition of, 5, 152, 386
necessity for, 5
normal, 384
other payments included in, 385
pure, 384
reasonable, 6, 154
relation to fixed and variable cost, 158
and risk, 386
true, 152-154
as wages of management, 384
Propensity to consume, 102, 433, 446-448
marginal, 447
Property taxes, 407
Proportional taxation, 401
Proportionality, principle of, 166-180
Proprietorship, 43
Psychic income, 346
Pump priming, 104
Pure competition, 139*n*.
(*See also* Competition)
Pure Food and Drugs Act, 444

Q

Quantity theory of money, 85

R

Rates, of exchange, 273-275
of taxation, 401-403
Reciprocal trade treaties, 273
Rediscount, 72
Rent, 320-335
contractual, 321
under cultivation, extensive, 327
intensive, 330-333
differences in, 323, 330
economic, 321
Rent, effect on price, 333
explicit, 321
formula for, 332
general level of, 323
implicit, 321
kinds of, 321
as a marginal product, 323
residential, 333
as a surplus, 322
Representative firms, 212
under actual competition, 214
and bulk-line price, 217
under dynamic conditions, 212-214
relation of average cost to long-run price, 213
Reproducible stock, 223
Republic Steel Corp., 250
Reserve banks, 67
Revenue, marginal, 174
Risk, relation to profits, 5, 386
Risk takers, 1, 2
Roosevelt, Franklin D., 423

S

Sales taxes, 407
Saving, and investment spending, 100
rate of, 102
relation to consumption, 100
and spending, choice between, 445
Sears, Roebuck and Co., 254
Services, definition of, 28
Short run, 140
monopoly price in, 224
Simultaneous determination, 313
Single proprietorship, 43
Smith, Adam, 16
Society, nature of, 18
Specialization, 286
Speculators, 91
Spending, rate of, 92
Standard, bimetallic, 52
gold, 51
paper, 54
single, 53

Standard Fruit and Steamship Co., 249
Standard of living, 296
Static economy, 189, 195, 206, 208
Stigler, George, 42
Stock, common, 45
  corporation, 45
  fixed, 222
  preferred, 45
  reproducible, 223
Stockholder as entrepreneur, 390-392
Substitutes, prices of, 263
Substitution, 264
Supply, changes in schedule of, 134, 137
  equilibrium with demand, 135
  factors affecting, 133
  fixed, 144, 156
  individual schedule of, 132
  during intermediate period, 183
  meaning of, 131
  static concept of, 309*n*.
Supply curve, for labor, 311
Swift and Co., Inc., 252

T

Taft-Hartley Act, 425
Tariffs, 274, 277, 406
Tax system, 397-399
Taxation, 295, 411
  and ability-to-pay principle, 399-401
  benefit principle, 399
  degressive, 402
  effect on incentives, 403
  and equal-sacrifice principle, 400
  excise, 406
  incidence of, 410
  kinds of taxes, 405-408
  limits to, 402-403
  local, 408
  national, 408
  for nonfiscal purposes, 404
  progressive, 402
  on property, 407
  proportional, 401
  rate structures for, 401-403
Taxation, regressive, 401
  sales, 407
  severance, 409
  shifting, 410
  special, 408
  state, 408
Temporary National Economic Committee (TNEC), 20, 22, 245, 248, 250, 253
Time preference, and savings, 372
Token money, 58, 58*n*.
Total product, 349
Trade, 273, 276, 406
  agencies fostering, 280
  attitudes of nations toward, 276
  balance of, 293
  basis for, 282
  and bills of exchange, 288, 291
  favorable, 291
  and gold standard, 286-291, 294
  international, advantages of, 276
    contrasted with domestic, 273, 279
  mechanics of, 287
  restrictions on, 295
  stake of U.S. in, 277
  visible and invisible items in, 292
Trade barriers, valid and invalid, arguments for, 295-297
Trade treaties, reciprocal, 273
Treasury notes of 1890, 59

U

U-shaped Curve, 166
Unemployment, 106, 108
Union, independent, 421
Union Carbide and Carbon Corp., 246
Unit costs, 297
United Fruit Co., 249
United Mine Workers, 416, 417, 419
United States dollar, 52, 289
  in foreign trade, 273, 277-279
  in monetary system, 52-55, 57-60
U.S. Department of Commerce, 99, 100
  national income, data on, 314

U.S. Rubber Co., 253
U.S. Steel Corp., 250
Universal Atlas Cement Co., 250
Utility, definition of, 28, 32
  diminishing, 13
  form, 29, 36
  place, 30, 36
  possession, 36
  service, 32, 36, 37
  time, 30, 36

V

Value, definition of, 31
Variable costs, 157
Variable proportions, principle of, 177, 178
Veblen, Thorstein, 438
Velocity of money, 82, 84, 99, 109
Visible items, 292
Voluntary cooperation, 19

W

Wages, 297, 341-360
  bargaining theory of, 355
Wages, exploitation theory of, 356
  fund theory of, 354
  general level of, 323
  and labor unions, 358
  and law of substitution, 350
  principles of, 341-343, 351
  as psychic income, 346
  rates versus annual income, 359
  real versus dollar, 358
  subsistence theory of, 353
  under competition, 347
  in U.S., 351
Wages and Hours Law (Fair Labor Standards Act of 1938), 423
Wagner Act (National Labor Relations Act), 423-425
Want gratification, principle of, 12
Wealth, definition of, 28, 29
Weighted aggregates, method of, 89
Wheeler-Lea Act, 444
Wilcox, Clair, 20
Wool Products Labeling Act, 444
World trade (*see* International trade)