

UNIVERSAL
LIBRARY

OU_154945

UNIVERSAL
LIBRARY

Call No. 55

Accession No. 1

Author G.P.

Carver, F. B.

Title Principles of Econo-

This book should be returned on or before the date last marked below

PRINCIPLES OF ECONOMICS

Revised Edition

By FREDERIC B. GARVER
UNIVERSITY OF MINNESOTA

and ALVIN HARVEY HANSEN
HARVARD UNIVERSITY



GINN AND COMPANY

BOSTON • NEW YORK • CHICAGO

LONDON • ATLANTA • DALLAS • COLUMBUS • SAN FRANCISCO

COPYRIGHT, 1937, BY FREDERIC B. GARVER AND ALVIN HARVEY HANSEN
COPYRIGHT, 1928, BY FREDERIC B. GARVER AND ALVIN HARVEY HANSEN
ALL RIGHTS RESERVED

346.6

The Athenaeum Press
GINN AND COMPANY · PROPRIETORS · BOSTON · U.S.A.

PREFACE

THIS textbook is intended for beginning students of economics at the college level. It places more emphasis on economic analysis than is perhaps usual in introductory treatises. The exposition stresses the combination of the factors of production, value and price, money and credit, business cycles, the functional distribution of income, the dynamic elements in distribution, and international trade and finance. The authors have given as thorough a discussion of theoretical principles as is deemed suitable for undergraduate students. They have also endeavored to make this discussion realistic and concrete at all points and to show the bearing of the principles of economics on the practical problems of both the social and the private economy.

This revision brings the text abreast of recent developments in economic theory. On the factual side it discusses the many important changes in economic institutions, legislation, and social control that have been introduced in recent years. The chapters on value (Book II) have been rewritten to conform to new developments in the analysis of price under conditions of partial monopoly or imperfect competition. The subject matter of these chapters has been rearranged so that the transitions from perfect competition to imperfect competition and from static equilibrium to dynamics are more clearly brought out. The chapter on industrial combination has been revised to bring into clearer relief the effects of restraints on competition upon price. In the chapter discussing public utilities the objectives of government regulation of utility rates and other prices have been set forth, as well as some of the more important difficulties encountered in the attempts that have thus far been made to regulate these rates and prices.

The chapters on monetary theory, price fluctuations, and business cycles have been modernized in the light of important recent research. The devaluation of the dollar, fiscal policy, monetary stabilization, the social-security program, and tariff bargaining and trade agreements are among the new topics considered in this revision.

The book retains the same general approach to economics as the first edition. Certain basic chapters, such as those on the combination of the agents of production, rent, interest, and the law of comparative advantage, are virtually unaltered.

The authors desire to acknowledge their indebtedness to all those who have offered suggestions for the improvement of this text. They are under particular obligation to the instructors in the course in Principles of Economics at the University of Minnesota for many suggestions, and to Professor Arthur M. Borak, Mr. Francis M. Boddy, and Mr. Edmund Nightingale for reading and criticizing certain chapters.

FREDERIC B. GARVER
ALVIN H. HANSEN

CONTENTS

BOOK I · PRODUCTION

	PAGE
CHAPTER I · Introduction	5
Economics as a science, 5 · Economic law, 9 · The subject matter of economics, 12	
CHAPTER II · Income, Consumption, and Production	16
Income produced and income consumed, 16 · Income utilization, 17 · Consumption and saving, 18 · Durable and nondurable goods, 19 · Production and human wants, 21 · Classification of productive activities, 23 · The source of utilities, 26 · Economic utilities, 26 · Production in the modern economic order, 27 · Control over utilities, 29 · Private income versus social income, 31	
CHAPTER III · The Regulation of Production	34
The individualistic order, 34 · The law of supply and demand, 34 · Assumptions of the individualistic philosophy, 35 · Public regulation, 36 · Unregulated private capitalism versus regulated capitalism, 37	
CHAPTER IV · The Organization of Production	42
The function of the entrepreneur, 42 · Forms of entrepreneurial organization, 45 · Corporations and the control of economic life, 50	
CHAPTER V · Combination of the Agents of Production	54
The background of social institutions, 54 · The factors of production, 55 · The relative scarcity of the factors, 56 · The law of diminishing returns, 57 · Agents of production, 60 · The best combination of the agents of production, 61 · Average product and marginal product, 63 · Cost per unit of output, 64 · Lowest unit cost for marginal entrepreneurs, 65 · Lowest unit cost for superior entrepreneurs, 66 · Large-scale production, 71	

BOOK II · VALUE

CHAPTER VI · Markets and the Marketing Function	81
Definition and classification of markets, 81 · The variability of price-making conditions, 83 · Marketing functions, 86 · Marketing and specialization, 89	
CHAPTER VII · The Price-Making Process : Demand and Supply	94
Price and value, 94 · The analysis of price, 95 · Market price, 96 · Demand and demand price, 97 · The individual-demand schedule, 99 · Demand curves,	

100 · Collective demand, 102 · Characteristics of demand schedules, 103 · Changes in demand, 109 · Supply and supply schedules, 110 · The determination of supply, 115 · Changes in supply, 116

CHAPTER VIII · The Price-Making Process : Market Price **118**

Reserved and unreserved supply, 118 · Price without reservation of supply, 118 · The significance of supply prices, 123 · Market price with reservation of supply, 124 · Effects of changes in demand; supply constant, 125 · Effects of changes in supply; demand constant, 126 · Effects of changes in both demand and supply, 126 · Marginal buyers and sellers, 126 · Equilibrium prices and actual prices, 128 · Equilibrium prices under conditions of monopoly, 128 · Summary on market price, 132

CHAPTER IX · Consumers' Demand and Utility **134**

The analysis of individual and collective demand, 134 · Utility, 136 · Utility and scarcity, 138 · The principle of diminishing utility, 139 · Marginal utility, 141 · Utility and the individual-demand schedule, 143 · Marginal utility and market price, 148 · Some practical applications of the marginal-utility theory, 149

CHAPTER X · Cost of Production and Normal Price **152**

How supply is determined, 152 · Cost of production and price, 153 · Normal price, 156 · The position of the individual seller under perfect competition, 158 · Prime and supplementary cost, 160

CHAPTER XI · Interdependent Prices **169**

The principal types of interdependent prices, 169 · Rival supply and price, 171 · Joint demand, 175 · Rival demand and price, 179 · Joint supply and joint costs, 181 · Overhead costs not joint costs, 182 · Summary of the theory of normal value, 184

CHAPTER XII · Equilibrium Price and Monopoly **186**

Complete monopoly and competition contrasted, 186 · Causes of monopoly, 189 · Monopoly and price discrimination, 190 · Monopolistic competition, 193 · Competition and manipulation of demand, 196

CHAPTER XIII · The Dynamics of Competitive Price **203**

The position of the individual producer under static competition, 203 · Entrepreneurs' differential costs, 204 · Cost differentials and imperfect competition, 209 · Improvements and differential costs, 211

CHAPTER XIV · The Dynamics of Price : Varying Cost **213**

Effect of changes in demand on cost, 213 · Constant cost, 213 · Conditions necessary for constant cost, 214 · Constant costs and the shifting of costs, 217 · Increasing cost, 218 · Entrepreneurs' differential cost and increasing cost, 221 · Forces counteracting increasing cost, 222 · Increasing cost and the incidence of new unit costs, 224 · The meaning of "decreasing cost," 227 · Declining cost is not decreasing cost, 233 · Limitations on decreasing cost, 234 · Decreasing cost and monopoly, 235

Contents

vii

PAGE
240

CHAPTER XV · Industrial Combination and Monopoly

The pool and the trust, 242 · The Sherman Antitrust Law, 243 · The "rule of reason," 244 · Efficiency versus unfair practices, 246 · Patents, 247 · Price leadership as the basis of control, 248 · Sharing the market, 252 · The trade association, 253 · The Federal Trade Commission Act and the Clayton Act, 254 · Recent trust policy, 255 · Summary and conclusions, 259

CHAPTER XVI · Public Regulation of Prices

262

The growing importance of public regulation, 262 · The nature of a public utility, 263 · The regulation of public-utility rates, 265 · Market value of utility property, 266 · Reproduction cost, 267 · The rate of return, 270 · Operating expenses, 271 · Municipal-utility prices, 273 · Railroad rates, 277 · Summary and conclusions, 281 · Agricultural prices, 281

CHAPTER XVII · Speculation and Prices

287

Production for the future, 287 · Methods of meeting business risks, 288 · Speculation and risk-taking, 290 · Effect of speculation on prices, 292 · The regulation of the exchanges, 294



BOOK III · MONEY AND PRICES

CHAPTER XVIII · Money and Credit

301

The functions of money, 301 · Money and circulating media, 301 · Monetary standards, 302 · Kinds of money in the United States, 303 · The standard of value, 306 · Deposit currency, 311 · Credit instruments, 312 · The volume of deposit currency, 313 · The Federal Reserve System, 314 · The monetary system, 317

CHAPTER XIX · Price Fluctuations

319

Relative prices, or index numbers, 319 · The making of index numbers, 321 · History of prices, 322 · Effect of price fluctuations, 323

CHAPTER XX · The Theory of Money and Prices

330

The value of money, 330 · The equation of exchange, 332 · The quantity theory, 335 · The income theory, 337 · Long-run price trends, 339 · Price movements since 1914, 341

CHAPTER XXI · Business Cycles

348

Industrial fluctuations, 348 · Various phases of the business cycle, 349 · Rainfall and business cycles, 355 · Self-generating cycles, 356 · Money and credit, 358 · Price fluctuations and the business cycle, 359 · Overproduction and underconsumption, 360 · Investment and the business cycle, 363 · The business cycle an unsolved problem, 365 · Desirability of elimination of the business cycle, 366

CHAPTER XXII · The Control of the Flow of Monetary Purchasing Power	368
Compensatory controls vs. industrial regimentation, 368 · Governmental budgetary policy as depression remedy, 368 · Banking policy and the cycle, 370 · Centralization of banking control, 372 · Qualitative versus quantitative control, 373	



BOOK IV · THE DISTRIBUTION OF WEALTH AND INCOME

PART I · *Functional Distribution*

CHAPTER XXIII · Wages	379
The meaning of distribution, 379 · The importance of distribution, 380 · The earnings of labor, 381 · Wages the price of labor, 382 · The marginal productivity of labor, 383 · The principle of substitution, 388 · The supply of labor, 390 · Decline in population growth, 394 · The subsistence theory of wages, 395 · Summary of the theory of wages, 397 · Imperfect competition and wages, 398 · Wages and improvements in technique, 401 · Principles of wage adjustments, 405	
CHAPTER XXIV · Wage Differences	407
Labor-income groups, 407 · The proximate causes of differences in wages, 410 · Fundamental causes of differences in wages, 411 · Differences in wages between different places, 413 · The wages of women, 417 · Minimum-wage legislation, 418	
CHAPTER XXV · The Rent of Land	420
Land as a factor in production, 420 · The meaning of rent, 421 · Method of determining rent, 422 · The demand for land, 423 · The amount of rent, 425 · Rent and differences in land, 427 · The generalized, or social, theory of rent, 430 · Agricultural rent, 432 · Urban rent, 433 · The unearned increment, 434	
CHAPTER XXVI · Interest	438
Nature and kinds of capital, 438 · Kinds of producers' capital, 440 · The return to capital goods, 441 · The superior productivity of capital, 443 · Time preference, 445 · The motives for saving, 447 · The supply of savings, 450 · The demand for savings, 452 · The capitalization of incomes, 455 · Types of interest, 457 · The loan market, 458 · Interest rates and the supply of money, 459	
CHAPTER XXVII · Profits	463
Business profits and economic profits, 463 · Pure profits as a share in distribution, 464 · Profits from the earnings of rare natural ability, 467 · Un-	

Contents

ix

PAGE

certainty in economic life, 468 · The entrepreneur as risk-taker, 470 · Risk-taking and profits, 471 · The real cost of risk-taking, 472 · Natural limitations on competition, 473 · Gains of monopoly and of exploitation, 474 · Summary and conclusion, 475

PART II · *Institutional Aspects of Distribution*

CHAPTER XXVIII · Trade Unionism and Labor Legislation 477

The origin of trade-unions, 477 · How trade-unions affect wages, 479 · The growth of unionism in the United States, 483 · Employers' associations, 486 · Trade-union demands, 486 · The methods of unionism, 480 · Collective bargaining, 489 · Unionism in the modern industrial order, 491 · The legal status of unionism, 493 · Government intervention in labor disputes, 495 · Minimum-wage legislation, 499 · Hours of work, 500 · Social insurance, 500 · Social Security Act of 1935, 501 · Economic principles of labor legislation, 506 · Panaceas and the labor problem, 508

CHAPTER XXIX · Some Dynamic Aspects of Distribution 510

The wage-earning class; wages, 510 · The effect of productivity on wages, 512 · The effect of price fluctuations on wages, 513 · The effect of immigration on wages, 518 · The effect of unionism on wages, 520 · The farming class; land value and land rent, 520 · The business class; profits, 527 · The investing class; interest rates, 532 · Other factors affecting the rate of interest, 536

CHAPTER XXX · Inequalities of Property-Ownership and Income 539

Distribution of wealth in Massachusetts, 540 · Distribution of wealth in two Minnesota counties, 546 · Distribution of income, 547 · Equalitarian distribution considered, 551 · Production versus distribution, 553 · Permanency of relative distribution of income, 553 · Property institutions and social utility, 557 · Functional distribution, 558 · Functional and personal distribution contrasted, 560

CHAPTER XXXI · Socialism 562

The extent of the socialist movement, 562 · Origin and development of socialism, 563 · Utopian socialism, 563 · Marxian socialism, 564 · Evolutionary socialism, 567 · Syndicalism, 568 · Guild socialism, 569 · Communism, 569 · Socialism and communism contrasted, 573 · Socialism and capitalism, 574 · Difficulties facing socialism, 574 · Waste in the present order, 576 · Value and distribution in a socialist order, 577

CHAPTER XXXII · Some Economic Aspects of Taxation 579

Government finance and economic principles, 579 · The nature of a tax, 581 · Nontax revenue, 581 · Kinds of taxes, 582 · The incidence of taxes, 583 · The capitalization of taxes, 587 · The amount of taxation, 588 · Increase in taxation, 590 · Government debt, 595 · Tests of a good tax system, 596 · The distribution of taxation, 597 · Progressive taxation versus proportional taxation, 600

BOOK V · INTERNATIONAL ECONOMIC RELATIONS

	PAGE
CHAPTER XXXIII · Foreign Exchange	607
<p>The meaning of foreign exchange, 607 · Forms of foreign exchange, 608 · Foreign-exchange practices, 608 · The rate of exchange, 611 · Pegging the rate of exchange, 615 · Exchange control, 616 · The international gold agreement of 1936 and exchange rates, 617</p>	
CHAPTER XXXIV · The Balance of Trade	619
<p>Our balance of trade before 1873, 619 · Our balance of trade as a debtor country from 1873 to 1914, 620 · Increased exports during the World War, 621 · The United States as a creditor nation, 621 · Economic effects of an unfavorable balance, so called, 625 · Factors governing the balance of trade, 627</p>	
CHAPTER XXXV · The Law of Comparative Advantage	629
<p>The basic reason for international trade, 629 · Comparative cost in regional trade, 630 · Comparative cost in international trade, 633 · The menace of foreign competition, 639</p>	
CHAPTER XXXVI · Tariff and Free-Trade Principles	644
<p>Tariff for revenue versus tariff for protection, 644 · Importance of our foreign trade, 647 · Arguments in support of the protective tariff, 650 · The validity of the arguments for protection examined, 652 · Industries injured by a tariff policy, 657 · British and American tariff policies, 659</p>	
CHAPTER XXXVII · The Trade Agreements Program	662
<p>Postwar American commercial policy, 662 · The Trade Agreements Act of 1934, 667 · The most-favored-nation policy, 674</p>	
INDEX	677

PRINCIPLES OF
ECONOMICS

BOOK ONE



PRODUCTION

CHAPTER I · Introduction



ECONOMICS AS A SCIENCE

Economics is a study of the price and value aspects of human activities and institutions. This very general definition is broad enough to cover every phase of economic life, and it is applicable to economic activity in a system of private property, state socialism, or communism. The methods by which man satisfies his wants depend upon the social organization under which he lives, the kinds of natural resources existing where he lives, and the technical knowledge of which he is master. But whatever the conditions, he has wants based upon fundamental biological needs and upon the culture patterns of his community. Productive activity applied to natural resources is essential in order to satisfy these wants. Even if man lived in complete isolation certain fundamental principles would apply to these activities. When he lives in a society and practices the exchange of goods, these elemental principles still govern; but in addition other factors enter, owing to the circumstance that he lives among other men and co-operates with them in wresting the means of satisfying his wants from natural resources.

It is important to note that the science of economics furnishes the tools for the analysis of any kind of economic order. In a competitive society operating on the basis of private property, the price system will function in a certain manner. In a monopolistic society operating on the basis of private property, it will function to produce considerably different results. In a socialistic society it will operate in a still different manner. The fundamental tools of analysis which the science of economics has developed are applicable to a study of any of these various types of economic systems. The trained economist is equipped with an assortment of "scientific apparatus," so to speak, to which additions are constantly being made, such as supply-and-demand curves, average and marginal revenue and cost curves, the principle of proportionality, constant and variable cost, and the like. At this juncture there would be no point in listing the whole arsenal of apparatus which the develop-

ment of the science has produced during the last two hundred years. But it is important to call attention to the fact that these tools of analysis are as important for the understanding and guidance of a socialist economic society as for the understanding and guidance of a capitalistic economic system.

At one time economics was said to be a purely deductive science. By this it was meant that the economist began any investigation by setting up certain assumptions, as in geometry, which were supposed to be axiomatic. There could, it was alleged, be no dispute about these assumptions. One of them was the axiom that every person tried to maximize satisfactions; another, that if the population kept on increasing, the average product per worker must ultimately decline—an inexact statement of the law of diminishing returns; still another was the tendency of population to outrun food supply; and, finally, it was assumed that production could be increased indefinitely by the aid of capital. Having set up his assumptions, the economist was supposed by deduction to work out all the possible effects of changes in population, supply of land, or demand. Reasoning and a knowledge of the assumptions were supposed to be sufficient.

As a matter of fact, the writers who made these statements about the nature of economic investigation never followed their own prescription. It must be evident to every intelligent person that human relations are not so simple as this methodology assumed. When these earlier economists seriously attacked a particular problem, they always relied to some extent on deductive reasoning; but they also attempted to find out from history and experience what had actually taken place when one of the changes under investigation had occurred.

The method of economics is the same as that of the natural sciences. Every investigation begins with a hypothesis. For example, the price of wheat has risen. If the inquiry is to find out what causes or, more properly, antecedent conditions brought this about, the first step is to single out certain possible causes. *Past experience* has shown that the price of a good may rise if there has been an appreciable decrease in the quantity produced and offered for sale. Here, then, is one hypothesis—one possible cause. If it were the only possible cause, then there would be no reason for investigation. We should know at once that whenever wheat rose in price, the supply must have fallen off. But there are a great many other possible causes. It may be that the demand for white bread has increased, and here too *past experience* tells us that this factor also may cause the price of wheat to rise. Again, when the

money offered for goods increases, without a concomitant increase in the goods, all prices are likely to rise more or less. This is another probability that must be investigated.

The economist confronted with the question "Why has the price of wheat risen?" sets up numerous hypotheses. Then he tries to get facts to ascertain whether any of the events that *might* have caused the price to rise have actually taken place. Similarly, a bacteriologist when asked to ascertain the "cause" of an epidemic must first know what sort of disease is at work. This is comparable to the known fact that the price of wheat has risen. Then by drawing on his knowledge—past experience of himself and others—he forms a certain hypothesis concerning the cause of the disease. By means of tests familiar to scientists working in the field of bacteriology he ascertains, if he can, whether the micro-organisms known to cause a similar disease are present. If he finds that they are and that there are no other malignant organisms that might cause the disease, his work usually is done. It remains for other scientists to name the remedy that will ameliorate the ravages of the disease or eradicate it altogether. And it remains for a still different group to persuade people to take the proper measures to ensure amelioration or cure of the malady.

Every piece of scientific investigation begins with some knowledge of facts, that is, observations, and some knowledge of the relation of facts. To discover new truth there is only one procedure. A hypothesis derived by deductive reasoning is set up, and then additional facts are gathered with which the correctness or incorrectness of the hypothesis is tested. When the investigation can be controlled completely, the method is that of the "controlled experiment." A certain object is put in such a position that only one thing—the cause under investigation—can produce an effect. This variable is then allowed or caused to impinge upon the controlled object, and the result is noted. For example, a container filled with gas is heated to a certain temperature, and the change in pressure is noted. From this experiment the effect of varying temperature on the pressure within the container can be measured. The cause is, of course, the heat applied. From this experiment, when carefully repeated, relatively exact relationships between temperatures and pressures can be established.

X There is no other method of discovering relationships in either the inorganic or the organic world. Often, however, individuals make guesses, or they arrive at conclusions by the process of deductive reasoning. But until these guesses and conclusions have been tested by experimental

evidence, they remain hypotheses. The statement just made applies to scientific investigations in economics as well as in the natural sciences.

In economics, as in all other studies that involve human relationships, controlled experiments are extremely difficult. Suppose, for example, that we wished to know the effect of changes in the price of bread upon its consumption. Assume further that it would be possible to secure the co-operation of the bakers in a small community. The price of bread is gradually raised to three times its customary level. The volume of sales is observed. Then the price is reduced to one third its customary level and the amount of sales again observed. The procedure is repeated several times and for different prices between these extreme limits. It might be supposed that these observations would enable us to set up a complete scale of the amounts of bread that would be purchased at different prices and the prices at which different amounts would sell. And this would be correct if we could be sure that the external conditions were always the same.

The results of such an experiment would differ, however, from those obtained in the laboratory. (1) We should not know to what extent the same results would be obtained if the experiment were duplicated in some other community. That they would not be vastly different we can safely conclude, but the divergence might still be so great that no mathematical "law of the demand for bread" could be set up from this experiment. It is otherwise, of course, in the sciences that have as their subjects inorganic matter. (2) If the price of bread should remain high for a long time it might, and probably would, happen that people would substitute other foods for it. If the use of substitutes became habitual, a lowering of the price would not bring about the same increase in sales as though no such habitual use of substitutes had been developed. Needless to say, an experiment with inorganic materials would be subject to no such variation. (3) If the experiment were repeated after the passage of a decade, we might find that the conditions affecting purchasing power or the habits of consumption had changed, and the results would be different. (4) Finally, although the results of the experiment could be considered relatively accurate for the group, they would not be accurate for any given individual included in the group. Each individual is governed in his actions by subjective factors and by external conditions. Since individuals are subjectively somewhat different and since no two live in exactly the same environment, there is always some deviation by individuals from the behavior pattern of the group of which they are members.

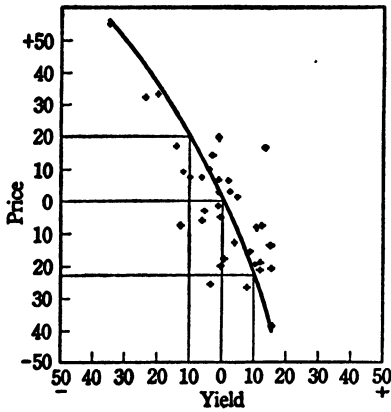
ECONOMIC LAW

It has just been said that economics does not pretend to predict the behavior of people as individuals. Nor does it pretend to predict group action with a high degree of accuracy. It does not set up any laws concerning the economic behavior of the individual or even of small groups. Speaking generally, it states that *on the average* and *in the mass* people behave in certain ways under the influence of certain stimuli. Similarly it asserts that if there are sufficient instances or mass data, relationships may be asserted. Thus, if sellers raise the price and hold it at the new high level, buying tends to decline. But this statement is true only if the buying is done by large numbers of people and if other conditions, such as the size of people's incomes, have not changed. Any individual may conceivably buy as much at the higher price. All economic laws are statements of tendencies. They are approximately correct statements of relationships.

Economic laws are tendencies rather than exact mathematical formulas. This follows, in part, from the fact that economic relationships are extremely complex. Many factors enter into every instance of economic change. While complex relationships hold for the phenomena of the natural universe also, the exact sciences are usually able to isolate causal factors in a definite and precise manner. Nevertheless it is because of the complexity of forces at work that no law of the physical universe ever appears to work out as stated in scientific laws. If the result observed in the actual world always corresponded with exact precision to the result predicted by science, it probably would not be necessary to engage in painstaking research to discover scientific laws. It is because of the complexity of factors at work that all scientific laws when first promulgated appear startling and unreal. Bodies do not actually fall according to the law of gravitation, for they are retarded by atmospheric pressure. The result indicated does not seem to occur in actual life. It is no wonder that so-called practical men have little faith in theory.

The conclusions of scientific study may properly be called laws if each set of conditions is subject to control so that each separate force can be measured and the resulting effect definitely predicted. In economics, on the other hand, many of the forces at work are often altogether unknown, others cannot be measured, and therefore the resulting effect of those forces which can be measured is subject to considerable variation, since it is vitiated by the unknown or the unmeasurable forces. Thus *exact* prediction becomes impossible, though prediction within

Fig. 1



The scattered points indicate the *actual* relationship between yield and price. The curve is fitted to these points and therefore represents the *average* relationship between these two variables. When the yield is low the price is high, and vice versa. The zero point, on the yield scale, represents a normal yield, and, on the price scale, a normal price. The figures on either side of zero represent deviations above (+) or below (-) normal. From the diagram it will be noted that if the yield is 10 per cent above normal, the price is about 20 per cent below normal; if the yield is 10 per cent below normal, the price is about 20 per cent above normal. (See Moore, *Generating Economic Cycles*, p. 36)

The Relationship between the Yield of Corn and the Price per Bushel

certain limits of variation often is possible. In such cases the term "tendency" is to be preferred to the term "law."

An illustration may help to clarify this matter. The accompanying chart (Fig. 1) shows the relationship between the yield of corn and the price per bushel (the trends having been eliminated in both cases). The curve represents the average relation between these two variables. If the yield is low, the price tends to be high; and if the yield is large, the price tends to be low. But the curve represents only an average condition, an approximation, a probability or tendency. It is impossible to say that with a certain yield the price will be exactly so much, because there are many other factors besides yield which affect price, and which have not been adequately taken account of. The curve shows what, approximately, may be expected in the aggregate, but the individual instance may be quite different. Moreover, the chart also shows the degree of variation that may be expected. Though the individual price may not conform to what on the average may be expected, we can be reasonably certain at least that it will not vary more than so much. That is a matter of no small importance.

In so far as economics can measure the probable result of certain conditions, together with the approximate variation from this result, it may properly be said to be a true science. In the main, economic laws

are of this statistical type. They hold for aggregates but not for individual instances. They set forth the average relationship between phenomena. They are approximations or tendencies.

We often hear people speak of the "natural" laws of economics, especially the law of supply and demand. We cannot alter, so we are told, the law of supply and demand. Now this statement indicates both a great truth and a great misconception. It contains a truth because it implies that a scientific law or tendency cannot be counteracted by blindly ignoring it—that the results follow from definite causes. It contains a great misconception, however, because it gives one the impression that human beings are helpless in the face of scientific laws. If certain conditions obtain, human beings are indeed helpless to avert results that follow, with a certain measurable variation, therefrom. But the conditions themselves may be changed; and when they are, then other results of a different sort will inevitably follow. In this respect economics and other social sciences do not differ from the natural sciences. Because certain conditions inevitably result in certain diseases, we need not despair and supinely submit to disease. We can alter the conditions and thus change the results. Economic conditions are especially subject to alteration because they are in part man-made institutions. Our economic life is therefore subject to human control; it is not all a blind play of uncontrollable external forces.

Economics is not and probably never will be an exact science in the sense that physics and chemistry are. It may well be asked, therefore, at the outset whether it is a science at all and whether any good can come from the study of it. To these questions satisfactory answers can be given. (1) Not all useful investigations can be made with the same degree of accuracy as those made in the laboratory of the chemist. (2) The worth-whileness of any study must be judged, not only by the accuracy of its results, but by the value of its results for social welfare. One might, to be sure, study economics for the pure enjoyment he might obtain from it, or merely to satisfy his curiosity about the behavior of certain economic phenomena. Such motives have surely acted as a spur to the study of astronomy, for example. But it is probable that more was done for the welfare of mankind by the discovery of the admittedly rough generalization called the law of diminishing returns than by all the exact calculations that have ever been made concerning the mass of the planets or the eclipses of the moon. It is evident that today many of the most serious problems confronting people are economic problems. Problems of technology—engineering, manufacturing—and of

health also are very significant. It would be futile to attempt to rank these problems in order of importance. But it is evident that unless the problems of money, wages, depressions, and prices shall come to be better understood, civilization may not have the means to carry forward the solutions of problems of technology and health, and people may not have the opportunity of enjoying the fruits of past discoveries in these other fields. The justification for the study of economics is, therefore, to be found in the belief that many useful explanations may come from that study and that solutions may be found for the many economic problems now confronting civilization.

THE SUBJECT MATTER OF ECONOMICS

The study of economics is assuming a place of increasing importance. The disturbed conditions following the World War and the unprecedented depression following the collapse of monetary systems in 1931 have enormously enhanced the interest of the general public in economic problems.

A similar situation existed a hundred years ago, after the Napoleonic wars; and the problems thrown up by that great struggle were in large part similar to those we face today. Sharp controversies arose over the causes of the conditions that prevailed and the proper remedies to be applied. British agriculture floundered in a tremendous depression, much like the one through which American agriculture has recently been passing. The landlord class demanded protection to shield it from the competition of the fertile land in the New World. British industry, on the other hand, wanted free trade that it might buy raw materials in the cheapest markets. A struggle arose between these two contending groups. Each had its advocates; each sought to prove that Great Britain would benefit by its own special program. Two great economists in particular stand out in this controversy: Thomas Robert Malthus, who defended the position of the landlord class, and David Ricardo, who pointedly set forth arguments favorable to the position of the industrialists. The struggle became a leading political issue and finally ended in a victory for free trade and the manufacturing class. In like manner tariffs and other foreign-trade obstacles are in the forefront of current discussion. But conditions have changed in many respects, and the outcome may well turn out to be utterly different from that reached a hundred years ago.

Another controversy that emerged in the post-Napoleonic period was

the currency question. England had temporarily abandoned the gold standard, and her paper money had depreciated. As a result there had been an enormous rise in prices. After the war these prices fell rapidly. Certain classes benefited from this; others suffered severely. What was the cause of the fluctuating value of money, and what could be done to stabilize the currency? What banking reforms were needed in order to prevent these violent fluctuations in prices? During the war England had borrowed huge sums of money, and the public debt had reached unheard-of proportions. How were the funds for meeting the interest charges to be raised? How rapidly might the debt be paid off? Who ought to bear the brunt of the taxes? These were typical economic questions dominant in the British politics of the time.

We today are faced with similar problems. Indeed, if one read certain parts of the *Principles of Political Economy* by John Stuart Mill and did not know that the book was published in 1848, one might well believe that the writer was grappling with the difficult problems confronting the world at the present moment. There are today, of course, many additional new problems or old ones presented in a different guise, some of which are an aftermath of the World War, others of the depression.

The war and the depression have brought about an increasing measure of governmental control over business and economic life. These controls have been directed in large measure toward a forcible readjustment of the interrelation of specific prices.

Now, as in the nineteenth century, the two great concerns in economic life are (1) to develop and foster such economic institutions and arrangements as will promote the highest possible standard of living and (2) to protect and safeguard the standard that has been achieved against serious setbacks. The first of these goals has to do with economic progress; the second, with economic security. So long as the standard of living was very low the former end was probably the more important. Currently, owing to the high degree of economic instability in the postwar world, and the consequent great difficulty of making secure the relatively high standard of living achieved, security and stability have become the paramount economic problem.

In part these current economic problems run in terms of certain economic groups. Thus in the United States a major problem both in the postwar decade and during the depression has been that of agriculture. In the depression agricultural gross income fell to about one third of the 1929 level. This was owing, not to a decline in agricultural output, but to a decline in the prices of agricultural products. This drastic

decline in gross farm income reduced net income still more, since certain expenditures could not easily be reduced, such as interest on farm indebtedness and taxes. The purchasing power of the drastically reduced net farm income declined by reason of the fact that the prices of the things farmers bought had not fallen as sharply as the prices of farm products. Moreover, the low purchasing power of farmers reacted unfavorably upon the market for industrial products and affected the prosperity of the urban community. Thus the agricultural problem ramifies into a whole network of general economic problems: the general fall of prices, the relation of agricultural prices to industrial prices, the burden of agricultural indebtedness, the problem of taxes, employment and purchasing power in urban communities, the cost of transportation, the export market for farm products, the value of the American dollar in terms of foreign currencies, and so on into a wide circle of inter-related problems. These problems led to governmental action in the form of the AAA, the Farm Mortgage Corporation, the Commodity Credit Corporation, the Farm Credit Administration, and the like.

A second major problem, especially during the depression, involved another large economic group, the urban wage-earners. The total wage income of urban workers also fell in the depression to about one third of the 1929 level. But this recession of income was very unevenly borne by urban wage-earners. The decline of total urban wage incomes was related in part to a reduction in wage rates, but much more to partial or total unemployment. A large fraction of the urban population suffered an almost complete cessation of income owing to the loss of employment and the inability to find other jobs. Measures such as the NRA and the Social Security Act were enacted, designed in part to deal with these conditions.

Other postwar and depression problems affect all economic groups and relate to serious defects in institutions which vitally concern the entire community. The collapse of the banking system in the United States in early 1933 forced upon the country the problem of banking reform. Similarly the losses that had been sustained by investors in corporate securities led to the passage of the Securities Act and the formation of the Securities and Exchange Commission. The special evils relating to the formation of a complicated network of intercorporate relations, particularly in the field of public utilities—layers of holding companies upon holding companies so that the underlying operating companies became all but submerged in speculative manipulations—led to the passage of the Public Utility Act of 1935.

These and other questions are being discussed on all sides—in the

newspapers and magazines, in hundreds and thousands of books, on the public platform, in legislative debates, and in private conversation. Every person in modern times is perforce more or less concerned with economics.

It must not be thought that any careful student of economics can give definitive answers to all these questions. There is a consensus among economists on a great many problems; many others are unsettled and there is no agreement concerning them. But even the unsettled issues will surely be approached more intelligently by those who have a firm grasp of economic principles. There are, moreover, many popular economic fallacies which can readily be exposed by any competent economist.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "A scientific law states that definite causes necessarily lead to definite results."—SELIGMAN. Comment.

2. "Everything that occurs in economic life takes place in accordance with some law; it is the function of the economist to ascertain that law."—SELIGMAN. Criticize.

3. "An economic law is a natural law so far as it states that given conditions will lead to given results. An economic law is not a natural law so far as it implies that human effort is impotent to modify the conditions which lead to the results."—SELIGMAN. Do you agree?

4. (1) The purpose of economics is to *explain* the phenomena of economic life. (2) The purpose of economics is to discover ways of attaining a larger measure of economic welfare. With which of these statements do you agree?

5. "We cannot in practice consider a fact without relating it to other facts, and the relation is a theory."—CLAY. Comment.

6. "The man who opens a discussion by saying that he is going to 'deal with facts not theories' does not mean that he is going to refrain from generalization; he usually means that he wishes to confine attention to a few facts that support his generalization, and to ignore all facts that conflict with it."—CLAY. Is this true?

7. Students often remark, "That may be true in theory, but not in practice." Is this possible? What is probably really meant by the statement?

8. What do you think of the following remark by a cynical critic of social reform: "Many people foolishly think they can substitute the laws of Congress for the laws of economics"?

9. "He [Herbert Spencer] taught me to look on all social institutions exactly as if they were plants or animals—things that could be observed, classified and explained, and the action of which could to some extent be foretold if one knew enough about them."—BEATRICE POTTER WEBB. Is this view sound?

CHAPTER II · Income, Consumption, and Production

INCOME PRODUCED AND INCOME CONSUMED

The total money income of the people of the United States was about \$80,000,000,000 in 1929. Owing to the fall in prices, output, wages, and employment, it fell to about \$40,000,000,000 in 1932. By 1936 the money income approached \$65,000,000,000. It is important to note that while the money income fell heavily because of the great decline in prices, the "real income," consisting of the quantity of goods and services, fell by a much smaller amount.

TABLE I. National Income¹ (in Millions of Dollars)

YEARS	MONEY INCOME	YEARS	MONEY INCOME
1929	81,034	1933	41,742
1930	67,917	1934	48,397
1931	53,584	1935	52,959
1932	39,545	1936	63,799

While there is a general correspondence in any period of time between the value of the commodities and services consumed and the value of the commodities and services produced, the two figures are not necessarily identical. In the usual case the value of goods² produced exceeds that of goods consumed. This is especially true in boom years. The excess of production over consumption is represented by the net accumulation of durable goods. All production, unless misdirected, results either in consumption or in the accumulation of durable goods.

¹The figures in this table are those for "income produced" as published by the National Bureau of Economic Research and the Bureau of Foreign and Domestic Commerce. (See Bulletin 59, May 4, 1936, of the National Bureau, and also the *Survey of Current Business*, July, 1936.) The definitions used are not altogether satisfactory and a revision of the methods used is in prospect. What is really wanted is estimates of income produced and of income consumed. The difference between these two would give net additions to and deductions from capital accumulations.

It should be noted that a considerable part of the income produced is not taken account of at all, such as the work of self-employed housewives and services performed by individuals for themselves or for their families.

²By "goods" is meant commodities and services.

These accumulations take the form of factories, machines, railroads, public-utility equipment, roads, drainage and irrigation works, stocks of finished and unfinished products stored in warehouses, automobiles, furniture, houses, and the like. In periods of depression the accumulation of durable goods may still be added to, but at a slower rate. In extraordinarily severe depressions, the existing stock of durable goods may not even be replaced. In these circumstances consumption exceeds production.

INCOME UTILIZATION

The manner in which income is utilized varies greatly according to the customs and habits of different individuals and families, but especially according to the amount of income received. In general, families with very low incomes are forced to spend a large percentage of the total on food. Families with large incomes spend a larger absolute sum on food, but this expenditure constitutes nevertheless a smaller part of the total. Such families, accordingly, have more left for cultural and recreational expenditures and for savings. This general tendency is clearly shown in the following table. It is derived from a composite

TABLE 2. Utilization of Family Incomes¹ (Urban Groups)

INCOME CLASS (IN DOLLARS)	AVERAGE EXPENDITURES (IN DOLLARS)					PERCENTAGE DISTRIBUTION OF EXPENDITURES				
	Food	Home	Attire	Other Living	Savings	Food	Home	Attire	Other Living	Sav- ings
0	360	250	110	110	— 60	46.7	32.5	14.3	14.3	
1,000-1,500	490	360	180	210	10	39.2	28.8	14.4	16.8	.8
1,500-2,000	610	470	250	310	100	35.0	27.0	14.4	17.8	5.7
2,000-2,500	700	580	300	470	180	31.4	26.0	13.4	21.1	8.1
2,500-3,000	770	690	350	640	280	28.2	25.3	12.8	23.4	10.3
3,000-3,500	820	820	390	800	400	25.4	25.4	12.1	24.8	12.4
3,500-4,000	860	960	450	960	510	23.0	25.7	12.0	25.7	13.6
4,000-4,500	900	1,110	490	1,120	620	21.2	26.2	11.5	26.4	14.6
4,500-5,000	920	1,250	530	1,260	780	19.4	26.4	11.2	26.6	16.4
5,000-6,000	940	1,460	590	1,480	990	17.2	26.7	10.8	27.1	18.1
6,000-7,000	970	1,720	680	1,770	1,320	15.0	26.6	10.5	27.4	20.4
7,000-8,000	1,000	1,930	760	2,050	1,720	13.4	25.9	10.2	27.5	23.0
8,000-9,000	1,030	2,100	820	2,280	2,240	12.2	24.8	9.7	26.9	26.4
9,000-10,000	1,060	2,290	900	2,500	2,720	11.2	24.2	9.5	26.4	28.7
10,000-15,000	1,150	2,620	1,040	2,980	4,270	9.5	21.7	8.6	24.7	35.4
15,000-20,000	1,300	3,600	1,400	4,300	6,600	7.5	20.9	8.1	25.0	38.4

¹ See Leven, Moulton, and Warburton, *America's Capacity to Consume*, p. 257.

of numerous sample studies. It presents an estimate of the general manner in which money is expended on food, housing, attire, other living expenses, and savings by urban families in the United States according to their place in the scale of incomes.

CONSUMPTION AND SAVING

The greater part of the national income flowing to laborers, farmers, investors, businessmen, and other producers is used to purchase goods for current consumption ; a part is saved. Usually the combined savings of many people are invested (either directly by the savers themselves or indirectly through banking or financial institutions) in capital or investment goods such as houses, factories, railroads, public utilities, or other durable goods. The flow of money income is thus expended on a roughly corresponding flow of commodities and services, partly consumers' goods, and partly producers' goods used by business in further production.

All producers compete for a share in the annual flow of money income. The competition of the sellers of consumption goods is patent to everyone. Advertising, personal solicitation, and all the techniques of directing the consumer's dollar into the treasuries of retail establishments we are accustomed to see going on around us every day. The makers of producers' goods likewise are competitors for a share in the flow of income. This competition is perhaps not quite so obvious as that of the retailers who sell directly goods for current consumption ; yet it exists. Such competition is evidenced by the reports of corporations in the financial pages of the daily papers, and in the advice given by forecasters and security market counselors as to the probable profitableness of alternative opportunities for investment. The company which erects an apartment house must compete with the promoter of a public-utility company for the savings of the community.

Thus consumers and savers are constantly exercising choices with respect to how they shall expend their incomes. Producers are constantly trying to influence these choices by advertising and salesmanship, and to a certain extent they control the choices made. But the choices are ultimately made by the recipients of the income. Every day they cast their ballots in the market place, and this stream of daily ballots becomes the guiding factor determining what shall be produced. Often the producers make mistakes. They forecast the poll of choices incorrectly and produce factories, automobiles, and textiles that do not

meet the demand of the market. Yet this mechanism ensures quick adaptation to changes in demand. Production is not cast in a fixed, frozen mold. It must constantly adapt itself to the choices and decisions of the income-receivers.

DURABLE AND NONDURABLE GOODS

A part of the income flow goes to purchase nondurable consumers' goods; a part goes into durable consumers' goods, such as houses, automobiles, and public works; and a part into producers' goods or instruments of production, such as machines and factory buildings. The division of the national income for the year 1929 into these categories is indicated in the rough estimates given in Table 3. From this table it appears that about 70 per cent of the total income was spent on nondurable consumers' goods and services, about 20 per cent on durable consumers' goods, while less than 10 per cent was expended on net additions to the stock of producers' goods. A part of the durable consumers' goods (probably about two thirds) is mere replacement, while a part represents net additions to the existing stock. If we add the net accumulation of durable consumers' goods to that of producers' goods we get an accumulation of new wealth equal to about 15 per cent of the total income. This represents the probable net savings for the year 1929. It must be remembered, however, that that year was the peak of the boom. Good and bad years averaged, it is probable that not more

TABLE 3. The Flow of Income into Durable and Nondurable Goods and Services¹

	1929 (IN BILLIONS OF DOLLARS)	PER CENT
1. Nondurable consumers' goods ² and services	56	70
2. Durable consumers' goods	17	21
3. Net additions to stock of producers' goods	7	9
4. Total national income	80	100
5. Replacement of producers' goods (Productions of these goods contributed indirectly to produce 1, 2, 3 above; hence this item is not included as a part of the total national income)	10	

¹ Estimates made by the authors in part derived from data in Bulletin 52, National Bureau of Economic Research, and in part from data in the *Survey of Current Business*, July, 1936.

² By nondurable consumers' goods is meant those commodities whose average useful life is less than three years.

than 12 or 13 per cent of the annual income represents net additions to the wealth of the community.

In addition to the division of the total income into the categories 1, 2, and 3 given in Table 3, an estimate is also made (item 5) of the expenditures made on the repairs and replacement of the existing stock of producers' goods. This estimate is placed at about \$10,000,000,000 for the year 1929. But this item is not included in the national income. If it were so included we should be guilty of double counting. The producers' goods which were serviced, repaired, and replaced were being used up in the process of producing the goods enumerated in items 1, 2, and 3.

Some workers are engaged directly in the making of consumption goods; others are so engaged indirectly in so far as they replace worn-out or obsolete productive equipment; still others are making additions to future productive capacity by building and constructing new factories and machines. When one examines occupational statistics, he is likely to be much impressed with the large numbers engaged in making goods which are not directly consumable. In so doing one is prone to forget that all workers who are preparing raw materials which enter into finished goods, or who replace, repair, and maintain equipment which is used in the production of these goods, are in a real, fundamental sense making consumers' goods. Superficially it does not appear so; yet in reality such is the case. Indeed, when account is taken of this fact it becomes clear that probably 85 per cent of the total national productive activity is devoted to supplying the needs of current consumption, while another 5 to 10 per cent of that activity adds to the stock of durable consumers' goods. This leaves some 5 to 10 per cent devoted to making new (additional) producers' goods.

As a first rough approximation, it may be assumed that the value of the capital accumulation (producers' goods and durable consumers' goods) is about five times the total annual income in the United States. If about 15 per cent of the income is saved annually—that is, invested in new durable consumers' goods and producers' goods—it follows that the capital accumulations of the country are being added to at approximately the average rate of 3 per cent per annum.¹

Properly speaking, the consumption in any one year is equal to the total flow of utilities or satisfactions derived from goods and services during that year. The utilities or satisfactions derived from an auto-

¹ Compare Gustav Cassel, *On Quantitative Thinking in Economics*. Oxford University Press, 1935.

mobile are enjoyed or "consumed" over a period of several years. So also with a house and other durable consumers' goods. Thus the volume of consumption may not have increased greatly in any given year, even though there was a large increase in the volume of consumers' goods produced in that year. The consumption in a given year plus the investment in new durable goods produced in that year equal the total income. And a part of the new investment may have been made in durable consumers' goods. Thus the volume of consumption and the annual output of consumers' (durable and nondurable) goods are not equal quantities in any given year. These are important distinctions; and if they are not firmly grasped early in the study of economics, much confusion is likely to result.

Durable goods, whether consumers' goods or producers' goods, are increasingly important under modern economic conditions. They are sometimes called capital goods or investment goods, because they involve an investment of capital. More and more our modern economic life is one which is profoundly affected, both on the side of production and on the side of consumption, by capital goods. Modern economic instability is, as we shall see later, largely related to the growing significance of capital goods. Thus the manner in which income is expended, the character of the community's consumption, and the flow of savings are very fundamental facts in the understanding of modern economic societies.

Producers' goods, that portion of investment goods which is used as instruments for further production, play an important role in modern communities. Primitive societies produced largely by direct methods, that is, by the direct application of labor to natural resources. Under modern complex conditions, production, for the most part, is carried on by an elaborate capitalistic process, by a roundabout system, in which expensive producers' goods are made first, in order that they may aid in the output of final consumers' goods.

PRODUCTION AND HUMAN WANTS

"Production," as the word is used in economics, means the utilization of labor and natural resources for the purpose of satisfying human wants. In the modern exchange economy, people are willing, either individually or collectively, to pay a price for production. But of course no one would be willing to do so did he not *anticipate* that production activities would result in the satisfaction of human wants.

These activities may give rise either to concrete tangible goods, such as the merchandise displayed in shops, or to direct personal services, such as the transportation of people by railroads, the work of physicians, entertainers, and government officials. The producers of *goods* and *services* are concerned primarily with getting a price for their activities, but the ultimate result of their efforts is nevertheless, in the usual and normal case, the satisfaction of human wants.

Some productive activities may not be desirable from the social point of view and in the opinion of moralists; but so long as someone is willing and able to pay for them, they are *economically* productive. Instances in which the moral and the economic criteria come into conflict are numerous. The manufacturer and the seller of adulterated goods or of patent-medicine nostrums are, in an economic sense, productive so long as they can find buyers. That their activities in an overwhelming majority of cases are socially pernicious most persons would concede at once. Hence laws are enacted to restrain these types of production. But since a price can be obtained for such wares, the makers and venders of them are engaged in production. The economist's position in this matter is not based upon a belief that economics has nothing to do with ethics. The effect of ethical standards upon all the affairs of life is too great to permit the consistent maintenance of such a belief. But if once one concedes that unethical acts are not productive, he has entered upon a course of reasoning which will lead to the conclusion that a very large part of human endeavor is unproductive from the point of view of different groups of people. The illustrations given are cases in which a majority would probably agree that production and sale were socially harmful. But what about the writing and sale of cheap, salacious fiction? or the production and exhibition of motion pictures that depict acts of violence or crimes? or the sale of whisky in countries where no legal prohibition exists? Certain groups of people in India have religious scruples against the use of beef for human food, but most of the world has not. Now economics strives to set up principles that shall be, as nearly as possible, generally applicable to all economic relations. The scope of its investigations cannot, therefore, be limited to occupations and acts that receive unanimous ethical approval. Moreover, since it deals primarily with the valuation of goods, it cannot refuse to consider valuations which concern commodities that harm rather than benefit mankind.

Economics is not, however, estopped from considering the effects of certain types of production or of consumption upon social welfare. If the consumption of alcoholic beverages affects the production of wealth

adversely, that is an economic fact and cannot be ignored. Or if industrial education in the public schools leads to greater efficiency in production, that fact must be given its proper weight in considering the reasons why a nation that provides such education can undersell another nation that does not. But economics considers these conditions only so far as they affect production, valuation, and distribution, and not as moral or educational problems.

Not all wants are satisfied through the activity of human beings; many are satisfied without any effort at all. The air we breathe satisfies an absolutely essential need, without a thought on our part. But a large range of human wants cannot be satisfied without effort. Even when we depend upon the forces and properties of nature, in large part these forces and properties must be adapted to our needs before our wants can be satisfied. Because most of our wants would otherwise go unsatisfied, we are willing to engage in productive effort.

CLASSIFICATION OF PRODUCTIVE ACTIVITIES

Productive activities may be classified under the following heads:

- | | |
|-------------------|---------------------------------------|
| 1. Agriculture | 6. Exchange |
| 2. Lumbering | <i>a.</i> Merchandising |
| 3. Mining | <i>b.</i> Banking |
| 4. Manufacturing | 7. Personal and professional services |
| 5. Transportation | |

The term "production" formerly was not used to include all the activities listed above. Even today the term is used popularly in a much narrower sense. Persons engaged in agriculture, lumbering, mining, and manufacturing are often spoken of as if they were the only "producers." When men produced exclusively for their family wants, or even later when they produced for a purely local market, there were indeed few other forms of productive activity. But with the widening of the market and with the coming of territorial specialization, when the people of one region devoted themselves to the production of one commodity and those of other regions to other commodities, a considerable amount of labor was of necessity employed in transportation and exchange. As the market became wider and the geographical specialization of labor increased, the producers of raw materials and manufactured products became increasingly separated from the consumers. This separation had somehow to be bridged by transportation, merchandising, and banking activities. Without these activities we should be forced back into the

primitive economy of small local markets. Such an economy would mean that we should lose all the gains of territorial specialization, since every region would be compelled to produce for itself all it needed. It would mean also that we should have to forgo all the advantages of large-scale production, since a small local market could not possibly absorb all the output of a large-scale plant. In place of the modern machine method of production we should have hand tools, small shops, and small-scale production. All this would entail an enormous loss of want-satisfying goods. Yet this is exactly where we should be were it not for the functions carried on by transportation and exchange. It must be evident therefore that labor engaged in transportation and exchange is a necessary part of a whole group of activities which results in yielding satisfactions and utilities. Such labor is therefore just as truly productive as labor involved in agriculture, mining, lumbering, or manufacturing. Any activity, other than play, which in any way contributes to the satisfaction of human wants is productive.

It is certainly true that satisfaction-rendering objects (goods) may be produced without the aid of middlemen, merchants, bankers, or even workers in transportation. They were so produced in large part during the handicraft age. But they could not be produced *in the manner* in which they are produced in modern times without the aid of these groups, and therefore not so cheaply. Transportation and exchange are vital and necessary parts of the modern method of production. The modern method of production is a complex chain of activities in which there are many links. Goods are no longer produced by one person or even by a special group of producers. Goods are produced by a society practicing extreme geographical specialization and division of labor. Any firms or persons constituting a link in that complex and specialized process are certainly engaged in the production of goods.

In the handicraft period the market, as we have already noted, was largely local. The exchange of products between crafts in the local market was a simple matter. The master craftsman was therefore worker and merchant, both in one. In the present economic order, with national and world-wide markets, the exchange of products between different regions, different industries, and different stages in the process of manufacture becomes an exceedingly complex matter. Each stage in the process is dependent on the preceding stage for the supply of raw material, and on the next stage for the market for its finished product. Each group of producers is dependent upon some other group of producers for a market for its products, and each group of producers in

turn constitutes the consumers of products manufactured by other groups of producers. These various groups of producers and consumers must be brought into contact with one another through the processes of transportation, exchange, wholesaling, and retailing.¹

It thus becomes evident that the country grain-elevator man, the warehouseman, the wholesaler, the retailer, and the banker are just as truly engaged in production as are the farmer, the miller, and the baker. Thus in the production of bread we have the following co-operating specialists: farmers, elevator men, railroad crews, manufacturers of railroad cars, coal and iron miners, iron and steel workers, lumbermen, wholesale merchants, millers, bankers, truckmen, bakers, retailers, and many others. The following tables give some illustrations of the relative importance of these various functions in the modern economic order.²

TABLE 4. Distribution of the Dollar Paid by Consumers for Bread

	1913	1921
Farmer	28.0	28.1
Transportation	2.8	2.6
Elevator	1.1	2.8
Flour-manufacturer	0.6	0.6
Transportation	2.4	4.4
Bread-manufacturer	43.6	42.9
Retailer	<u>21.5</u>	<u>18.6</u>
	100.0	100.0

TABLE 5. Distribution of the Dollar Paid by Consumers for Men's Shoes

	1913	1921
Raw materials	40.8	38.5
Manufacturer	30.5	33.2
Retailer	<u>28.7</u>	<u>28.3</u>
	100.0	100.0

**TABLE 6. Growers' and Distributors' Returns on California
Oranges per Box**

	FIVE-YEAR AVERAGE 1913-1918	DECEMBER 1, 1920, TO NO- VEMBER 15, 1921
Growers received for fruit on tree	\$2.27	\$2.30
Harvesting, packing, and selling	0.51	0.88
Transportation	0.94	1.62
Wholesalers	0.48	0.61
Retailer	<u>1.50</u>	<u>2.06</u>
Consumer paid	\$5.70	\$7.47

¹ For a fuller discussion of the marketing function in production see Chapter VI of this book.

² See Report of the Joint Commission of Agricultural Inquiry, Part IV, 1922.

Early writers on economics thought of persons rendering personal and professional services as unproductive laborers. A little reflection, however, will soon convince one that it is very illogical to call the maker of a surgical instrument a producer and refuse to admit that the surgeon who uses the instrument is a producer. A violin in the hands of a great artist satisfies wants. The flow of utilities thus rendered is produced both by the maker of the violin and by the artist who plays it. Both are producers, for their activities directly or indirectly render satisfactions, utilities, services. All activity that results in the satisfaction of wants, and for which someone will pay a price, is productive.

THE SOURCE OF UTILITIES

The satisfactions, utilities, or services that come to us have various sources. Many of our wants are satisfied by nature directly, without any activity of man at all. Illustrations at once suggest themselves: the air we breathe, the heat and light from the sun, the running brook, the fruit on the trees, the beauties of rocks and mountains, of valleys and the setting sun. Other wants are satisfied by the action of man engaged in fashioning, shaping, or transforming the objects of nature in such a way that they will be capable of yielding satisfactions. Thus materials are taken from mine and forest and fashioned into houses, furniture, and musical instruments. In his labor man is constantly utilizing forces of nature which assist him in his work. The forces of waterfall and wind, heat and electricity, are harnessed to assist him in his work. Other wants are satisfied directly by the activity of human beings. As examples we may cite the utilities rendered by servants, barbers, preachers, teachers, singers, and actors.

ECONOMIC UTILITIES

Some of the utilities referred to above are rendered by nature in such abundance that the wants to which they minister are fully satisfied. For example, air exists in such quantity that all want for it is completely satisfied. But most of the want-satisfying services rendered by nature and man are scarce. They are not offered in sufficient abundance to satisfy all wants. If one unit of these services were lost, a certain amount of human satisfaction would be lost. Each unit matters; it has significance, importance, worth. Because of this we act economically toward those services which are scarce. We act on the assumption

that there is not enough of them to go around. We economize. We therefore call them *economic* utilities, services, or satisfactions.

But we should not act economically toward them if we were able in no way to *control* the utilities, services, and satisfactions rendered. In a northern climate the heat from the rays of the sun is not sufficient to satisfy human wants, but the utilities derived from them nevertheless are not economic utilities because nobody has so far been able to control in any way the rendering of these utilities. Since we cannot control the rendering of these utilities, it is impossible for us to act economically toward them; and they are therefore not economic utilities. We are concerned in our study particularly with those utilities, satisfactions, and services which are *scarce*, which we therefore wish to act economically toward, and which we *can act economically toward because we are able to control them*. The importance of this question of control will become apparent later when we consider the subject of property and income.

It sometimes has been asserted that the test as to whether utilities, satisfactions, or services are economic or not depends upon whether or not human labor has been applied to them. It must be obvious from the foregoing discussion that this is not the case. Many utilities derived directly from nature without any assistance on the part of man are economic. The utility afforded by the surface of the earth as a place to stand upon and to build upon is an example; likewise the utilities rendered by lakes and rivers, wild fruit trees, and wild animals. Economic utilities are not necessarily created either in whole or in part by labor. They are derived in part from nature, in part from natural objects fashioned and transformed by man, and in part from man directly.

PRODUCTION IN THE MODERN ECONOMIC ORDER

It is now clear that the flow of utilities, services, and satisfactions which constitutes our real national income comes in part from nature and in part from the productive services of man. We shall now attempt to analyze the way in which productive services are rendered in the modern economic order. To make this analysis clear let us briefly contrast the modern method of production with the method of the handicraft period. In the handicraft period there was little specialization or division of labor within the industry. The master craftsman or the journeyman performed all the various processes involved. There were no complex and highly specialized agents to be co-ordinated, as is true

under the present system. Contrast this method of production with the modern process of making shoes, for example. Here we have a great variety of highly specialized machines, and every step in the process requires a different kind of labor. These minutely specialized processes must be properly co-ordinated. Foremen, managers, and superintendents are needed. But this is not sufficient. Not only must we have many types of manual laborers, clerks, technical experts, accountants, managers, and superintendents, but also some person or group of persons must promote the business in the first place and later assume the ultimate control and responsibility for successfully carrying the venture on. There must be "captains of industry," or *entrepreneurs*.¹

In the modern economy it is not always easy to locate the *persons* who perform the function of entrepreneurship. In small-scale businesses the individual proprietors are of course the entrepreneurs; but in large-scale businesses the function is performed by a business organization, such as a corporation, or, rather, it is performed by persons functioning through such an organization.

It must not be assumed, however, from what is said above that in the primitive local economy the entrepreneurship function was lacking. But it is certainly true in the modern economy, with its complex system of large-scale production, that the function of entrepreneurship is far more important than was the case in the days of local self-sufficiency. The modern entrepreneur stands at the center of the economic organization. All the specialized agents of production are under his control. It is he who sets the productive resources in motion.

In the next place it may be observed that in the handicraft period production was carried on by a short-time direct process. The tools used were suitable devices to assist the work of the hands. Little work was involved in making these tools compared with the work employed in making the finished product. In the modern system production is carried on by a roundabout, time-consuming process. Much labor is involved in the production of machines, factories, railroads, and other producers' goods. When machinery has been built, it is offered for sale by the manufacturing company. From the standpoint of economic theory we may conceive of this machinery as the source of uses which flow from it through a succession of years. These uses will assist in the production of consumption goods which will finally satisfy human wants. It is because the machinery will indirectly but ultimately aid in satisfy-

¹For a fuller discussion of the function of entrepreneurship see Chapter IV of this book.

ing human wants that it has value. It may therefore be thought of as a source of future utilities. But no one prizes future utilities as highly as he does present utilities. Therefore the present value of the machinery must be considerably less than the sum of the values of the utilities it ultimately will render. Now if the manufacturing company is to be able to sell this machinery, someone must be willing to give up present purchasing power in exchange for it. Indeed the machinery would never have been built, had not the manufacturing company been fairly certain that someone would be willing to give up present purchasing power for this source of future utilities. But no one would do so were it not that the sum of the future utilities which will finally be received will probably be more than the utilities represented by the purchasing power given up. The person who gives up present purchasing power in exchange for the machinery is to that extent forgoing immediate satisfaction of wants. Since capital equipment would never be produced unless someone was willing and able to buy it and hold it until its future utilities were forthcoming, and since no one can buy it without giving up present satisfactions, it is obvious that under the modern system of production, carried on as it is by means of elaborate capital equipment, this service of waiting, or forgoing present goods, which takes the form of saving, is an absolute necessity.

The services rendered under the modern system of production may then be summarized as follows:

1. The services of labor
 - a. Manual labor
 - b. Mental and technical labor
2. The services of entrepreneurship
3. The services of waiting, saving
4. The services of nature (natural resources)

CONTROL OVER UTILITIES

We have so far considered the source of the continuous flow of utilities, satisfactions, or services. We shall now consider the means whereby in modern society control is exercised over this flow of utilities and services. So far as the utilities are subject to that sort of control which can be bought and sold, which can be exchanged in the market, we have the institution of property. Property, the court says, is anything which has exchange value, and it would not have value if it did not give control over utilities. Property is therefore the legal claim which gives transferable control over utilities and services. So far as utilities

are not subject to that sort of control which is transferable, the control is exercised through the institution of personal freedom. Since slavery was abolished, the law has not permitted one person to transfer to another permanent control over the utilities his person is capable of rendering.

A portion of personal and professional services is controlled directly by the persons rendering them. The services rendered by the servant form the clearest illustration. But the proportion of utilities thus received by consumers is very small. Take the case, for example, of the motion-picture theater. The utilities rendered in this instance are a product of a whole network of contributing factors. There is the theater building with all its furnishings, and the motion-picture machine with all the labor, saving, entrepreneurship, and natural factors employed in its production. There is the film with all the actors and other persons who helped to produce it. Whole industries assist in the process of furnishing motion pictures to the public. The actors do not sell motion-picture utilities to the public. The utilities rendered are under the control of the entrepreneur, or rather a group of entrepreneurs.

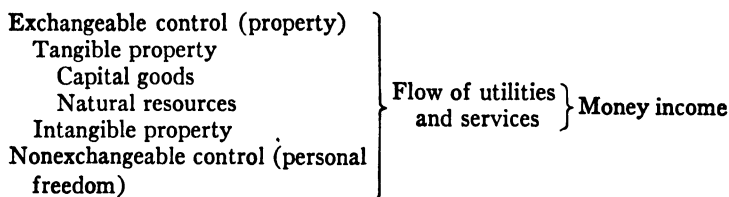
Control implies the power of *withholding* these utilities, services, and satisfactions. The power to withhold gives the bargaining power to demand services in return for services rendered. The flow of utilities, satisfactions, and services derived from this control constitutes income. This exchange of services against services takes place on the basis of the importance of each. The significance of each service, utility, or satisfaction is measured in terms of other utilities, services, or satisfactions. The establishment of the terms of exchange of one type of service for another is the process of valuation. Thus the value of property arises from the flow of utilities and services which may be obtained in exchange for the services over which it gives control.

Now our motion-picture entrepreneurs have the power to yield or withhold the utilities of motion pictures. They are therefore able to command a flow of utilities in exchange for them or, what amounts to the same thing, to assert a claim (in the form of money income) upon a flow of utilities. True, they are compelled to share this claim upon a flow of utilities with all the contributing factors employed by them; the laborers, the savers, the owners of natural resources. If competition has been at work long enough to have attained its full influence, the entrepreneurs will be compelled to share all the money income with the contributing factors, reserving only enough for themselves to compensate for their own contribution. But competition may not be at all effective.

The motion-picture entrepreneurs may be sheltered from competition by patents, secret processes, combination, or good will, so that they will not be compelled to hand over all the money income to the contributing factors. Such control gives rise to a kind of property which we designate "intangible property."

The foregoing discussion may be summarized as follows:

1. Property is the institution which gives transferable control over utilities and services.
2. Personal freedom is the institution which gives nontransferable control over the utilities and services which may be rendered by one's own person.
3. Income is the flow of utilities or services which one can command because of control over other utilities and services.
4. Money income represents a claim upon the flow of utilities or services which one can command because of control over other utilities and services.
5. Intangible property represents transferable control over utilities or services rendered under conditions which permit of *unusual* withholding power, such as may be derived from patents, secret processes, good will, or combinations. The excessive power thus derived to withhold the utilities or services means excessive power to demand services in exchange for services rendered. This excessive power is represented by "intangible assets."
6. The following outline indicates the interrelations mentioned above:



PRIVATE INCOME VERSUS SOCIAL INCOME

Patents, secret processes, and monopoly control are illustrations of intangible-property rights. They give the owner the power to limit output and thus to secure a *larger price* for a product made scarcer by these restrictive devices. It may even be that by the use of these methods he is able to get a *larger income* from a smaller output of goods. Is he then more or less productive? From his own individual point of view the smaller production yields the larger income, but from the standpoint of society as a whole he is less productive because he produces a smaller quantity of want-satisfying goods. It is true that in the modern exchange economy no act is productive, from an economic standpoint, which does not command a price; but we cannot measure the *extent*

to which an act is productive, from the social point of view, by the price that its products are able to fetch in the market. While the product in question is enhanced in value by restriction of output, the value of other things is thereby lowered, the total volume of production is lessened, and the total real income of society is reduced. The final test of the *degree* to which an act is productive is the volume of economic goods and utilities resulting therefrom.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Material goods and personal services which render utilities that are *scarce* are referred to in economic literature as *economic goods*. How do you distinguish between economic goods and free goods?

2. Economic goods may be defined as goods which (1) yield utility, (2) are scarce, (3) are subject to human control. Defend this statement.

3. The term "property" is sometimes said to be synonymous with the term "economic goods." Criticize.

4. "Property" is sometimes defined as "material economic goods." Criticize.

5. The early writers on economics held the view that only such laborers as turned out material goods were productive. Adam Smith placed in the class of "unproductive" laborers "some both of the gravest and most important and some of the most frivolous professions; churchmen, lawyers, physicians, men of letters of all kinds; players, buffoons, musicians, opera singers, opera dancers." Is this a useful distinction?

6. In the terminology of the business world economic activities are divided into production, marketing, transportation, finance, and so on. In economic terminology the word "production" includes all these activities. There is something to be said in favor of both uses. Discuss.

7. "The view that production means the creation of satisfactions or utilities is naive and old-fashioned; production in point of fact means doing anything for which people are willing to pay a price." Criticize.

8. "The real test as to whether labor is productive is this: Does it earn an income? If so, it is productive." Do you agree? Is the purveyor of quack medicines productive? the professional gambler? the seller of worthless mining stock?

9. Is there a sound distinction between the terms "criminal labor," "predatory labor," "acquisitive labor," and "productive labor"?

10. Productive labor, in the business sense, means labor which yields a money income; productive labor from the standpoint of human welfare means labor which has social utility. Which position should the economist take? Why?

11. A monopolistic combination restricts production and so raises the price. As a result it increases the total value of its products, but decreases the physical quantity of its products. Is it more or less productive? Explain.

12. "As has been intimated more than once in the course of this discussion, the limitation of supply to enhance the price is a normal and legitimate business practice by business men. . . . But modern trade-unionism partakes of the nature of business enterprise which implies that its members are more interested in wages—price per hour, per day, per week—than in output of goods—tons and yards per day, per week, per month."—ARDZOONI. If this is so, are trade-unions productive?

13. The statistician Engel made a study of the consumption of the working and lower middle classes in Saxony in 1857. He found that as incomes increased, relatively less (in terms of per cent of the income) was spent on food, more on clothing, more on education, health, and recreation, while the per cent of income spent on lodging, light, and fuel remained constant. Compare with the table given in this chapter.

CHAPTER III · The Regulation of Production



THE INDIVIDUALISTIC ORDER

Adam Smith, who wrote in 1776, just at the beginning of the Industrial Revolution, was an ardent advocate of *laissez faire*. By that he meant an economic order in which there is a minimum of governmental regulation of industry and in which each individual is left free to engage in any business or calling that he may wish, to accumulate and use private property without undue restrictions by government, to enter freely into such contracts as he may desire without restrictions by law, to exchange products on the market at such prices as he can get. There are therefore four fundamental institutions which form the basis of a *laissez-faire* or individualistic economic order: (1) freedom of enterprise, (2) private property, (3) freedom of contract, and (4) freedom of exchange.

Such an economic order is sometimes spoken of as an unregulated society, but in fact this is not the case. The institutions of private property, freedom of enterprise, freedom of contract, and freedom of exchange constitute a type of social control of economic life. None of these institutions can exist without some form of social organization to give them effect, and all are products of a long process of social evolution and economic development. All have important regulative effects upon our economic activity. A society, however, which is subject to no sort of regulation, if indeed such a society could exist, is completely anarchistic. The institutions of private property, free contract, freedom of exchange, and freedom of enterprise are maintained and enforced by the sovereign power of the state and are themselves great regulative forces in the economic world.

THE LAW OF SUPPLY AND DEMAND

An individualistic order based on the four institutions named is often said to be regulated by "natural" or "automatic" forces, in contradistinction to the regulation imposed by conscious and deliberate social control. Thus in a society with private property, free contract, freedom

of enterprise, and complete freedom of exchange there is no governmental body deliberately determining how many bushels of wheat or how many pairs of shoes shall be produced or the price at which these goods shall be sold. These matters are settled, it is said, by the natural and automatic working of the law of supply and demand. If there is a shortage in the production of wheat, the price will be exceptionally high, and farmers thereby will be encouraged to produce more wheat. If plumbers are scarce, the wages of this class of workmen will be higher than for others of similar skill, and consequently workmen will be attracted to this field. Under the automatic operation of the law of supply and demand market price becomes the guide, indicating how much of each thing ought to be produced and how the working population should be distributed over the field of industry.

This regulative principle is, however, only a long-run tendency. It assumes the swift mobility of labor and capital. But capital invested in a certain plant cannot readily flow into other more remunerative fields. Farmers cannot instantly change from one crop to another; nor can they easily transfer their capital from farming to manufacturing when they find that the prices of farm products are unduly low compared with the prices of manufactured goods, as has been true during the recent agricultural depression. Nor can wage-earners in other fields readily turn plasterers when plasterers' wages are phenomenally high, as was true in some cities a few years ago.

ASSUMPTIONS OF THE INDIVIDUALISTIC PHILOSOPHY

It is assumed by those who have complete faith in the automatic working of the law of supply and demand that market price always can be relied upon to indicate what ought to be produced. But as the French economist Gide has well said, the law of supply and demand distributes production, not according to man's true needs, but according to his desire and the price he is willing and able to pay. Now there are many moralists who are ready to assert that people buy ugly things when they ought to buy things which are artistic and beautiful; that they spend too much money on automobiles, motion pictures, confectionery, and cosmetics and not enough on education; that the rich, through their purchases, direct production to luxuries which have very little social utility, whereas the productive power of the nation might be spent much more wisely in supplying things that would raise the standard of living and the material well-being of the masses. It is maintained that the law of supply and

demand is a very imperfect guide to production for two reasons: first, because income is unequally distributed, and so luxuries are produced for the few even though the necessities and comforts of life are sadly lacking for the many; second, because all buyers, both rich and poor, have wants and desires that do not conform to the canons of the moralists. Hence it is asserted that a more equitable distribution of wealth and income would result in a more rational demand for goods, and that education, culture, and morality would raise the standard of consumption.

The individualistic philosophy assumes that everyone will follow his own self-interest in a rational way and that if everyone does follow his own self-interest, the greatest social good will result. But neither assumption is wholly tenable. People are too ignorant and easygoing to follow their true self-interest. Unless restrained by law they buy adulterated foods, worthless mining stock, etc. Workmen do not refuse to work in plants where the machinery is improperly safeguarded. People are induced by expensive advertising and high-pressure salesmen to buy things they do not really want. Nor does the self-interest of each lead to the good of all, particularly in a society of unequals. The strong and clever are all too likely to exploit the weak and ignorant.

The so-called natural or automatic working of the law of supply and demand implies, moreover, perfectly free competition. It assumes that there will be no artificial restriction of supply by means of monopolistic combinations. It assumes that if prices are unusually high in a certain field of industry, the large profits resulting therefrom will attract more capital and labor, production will increase, and the price will come down to a fair level in line with the prices of other commodities. Thus it is expected that competition, and not governmental regulation, will operate to ensure a fair price. But in point of fact it is as natural to combine as it is to compete. And so we have our trade associations, monopolistic combinations, trade-unions, farmers' organizations, professional associations, and the like. Self-interest impels competitors to combine.

PUBLIC REGULATION

Modern societies have all reached the point where automatic regulation of the economic order breaks down. When the stage of combination is reached, the public in self-defense can no longer allow economic affairs to run their own course. *Laissez faire* no longer gives adequate protection to the public interest. The public, through the state, must take

hold and apply conscious and deliberate regulative measures. This may be done in one of two ways: first, measures may be adopted which aim to maintain competition and prevent combination; second, the combination may be accepted as inevitable, and measures may be adopted to control the combination in the interest of the public welfare. One such measure is the regulation of prices. It is intended not merely to secure a "fair" price but also adequate production. Prices may be regulated directly, as is done, for example, in the case of railroad rates; or they may be regulated indirectly by watchful oversight, as is done in the packing and steel industries. Another measure by which regulation could be applied would be by taxation of the surplus incomes accruing from combination. While this method might insure a fair profit, it would not insure a fair price or a proper amount of production, because monopoly prices necessarily involve restriction of output.

UNREGULATED PRIVATE CAPITALISM VERSUS REGULATED CAPITALISM

Adam Smith and the early classical economists believed that the automatic regulative factors of private property, free contract, freedom of enterprise, and free exchange were sufficient to insure economic welfare. Such has not proved to be the case. No doubt these regulative principles would work out satisfactorily if human beings had reasonable equality of bargaining power, and the knowledge and intelligence to see the various alternative opportunities, and the economic independence to withhold their services and goods with a view to making use of the best available opportunity. But human beings are grossly unequal in knowledge and power, and the inevitable result is the oppression of the weak by the strong. For this reason England found it necessary early in the nineteenth century to abandon the policy of relying exclusively on the automatic principles enumerated above, the policy of *laissez faire*, the policy of a minimum of governmental interference in industrial affairs. Individual freedom of contract was restricted by governmental interference in the interest of the public welfare. In 1802 a law was passed prohibiting the employment of apprenticed children under nine years of age in factories and limiting the hours of labor to twelve for apprentices under sixteen years of age. This act applied chiefly to pauper children who had been apprenticed to millowners by the parishes whose charges they were. With the coming of steam power the mills could be located in populous districts, and it was no longer necessary to

transport children (chiefly apprentices) to the remote places where water power was available. Hence apprenticed children were decreasing in importance and children living at home became of greater importance as factory labor. Accordingly, in 1819 the provisions of the earlier act were extended to all children working in the cotton mills, including those who lived at home and who were not apprenticed. Both these acts were left to local enforcement and were therefore largely ineffective. In 1833 an act was passed applying to all textile establishments except silk mills. This act completely forbade the employment of children under nine and limited the hours of labor for children under thirteen to nine a day and forty-eight a week. The act prescribed two hours of schooling a day for child laborers. For the first time reasonably effective enforcement was made possible by the appointment of factory inspectors. In 1842, after a Parliamentary investigation revealing shocking conditions in the employment of women and children in the coal mines, an act was passed which prohibited the employment of women and children underground. In 1847 the Ten-Hours Law, restricting the hours of labor for women and children in textile mills to fifty-eight a week, was passed. These early acts themselves bear witness that the laissez-faire policy resulted in the exploitation of the weak by the strong and failed to secure the nation's health, morals, and general welfare.

The policy of social control has been carried much farther in England during the last eighty-five years. The act of 1864 contained provisions for sanitation. The act of 1867 extended the existing laws to include practically all kinds of factories and workshops. In 1901 the minimum age for child workers was raised to twelve years. In 1909 a minimum-wage law was passed by which wages in a few low-paid industries were fixed by Trade Boards. The scope of this act has been greatly extended, in 1912 being made to apply to coal mines, and in 1918 being extended to any industry where labor is not well organized and wages are relatively low. This act now covers about three million workers, largely women. Legislation was passed in 1897 by which employers are compelled to compensate workmen for industrial accidents. In 1911 compulsory insurance against sickness and unemployment was introduced and the employer was compelled to carry part of the expense burden. In 1909 the Housing and Town-Planning Act was passed, which enabled local public authorities to tear down unsanitary tenements and to purchase land and erect clean and wholesome dwellings. After the World War an extensive housing program was undertaken, and unemployment

insurance was extended to include practically all industries. This review of British social legislation is by no means complete, but is sufficient to illustrate the general tendency.

Similarly in the United States the automatic regulative forces operating by themselves alone have been found wanting. From the very beginning it was found necessary to regulate the banking business. In this matter free competition, freedom of enterprise, and self-interest were not sufficient. We have found it necessary to adopt a national program for the conservation of our natural resources. Freedom of enterprise and competition proved wasteful. We no longer permit immigration to be controlled by automatic forces, but definitely restrict the numbers that may come in from Europe, Australia, Africa, and Asiatic Russia, and exclude immigrants from the rest of Asia. We have found it necessary to prevent unfair competition and trade practices, to dissolve trusts, and to regulate the railroads and public utilities. Market price as a guide to the production of educational services was found wanting, and so we have made education free to all, including in many places books and sometimes meals for school children. City zoning laws substitute conscious city-planning for the blind play of automatic forces. Employers are compelled to compensate workmen for industrial accidents; most of the states have passed unemployment-insurance laws; and the Federal government has adopted a comprehensive national old-age-insurance measure.

But while we have departed widely from the principles of *laissez faire* and are moving steadily in the direction of more and more social control of industry, it is nevertheless true that the automatic forces are still of great importance. They are the all-pervasive factors that regulate millions of contractual relations and of purchases and sales in our economic order. Governmental control is important, but it does not enter every nook and crevice of our industrial life, as do the automatic regulative forces. Suppose we should attempt to abolish freedom of trade, private enterprise, and competition, and to substitute state control over each one of the millions of transactions that take place year in and year out in our complex industrial society. The task would be stupendous. One has only to contemplate the magnitude of such an undertaking to realize the extent to which the automatic regulative principles control our modern economic life.

These principles themselves may be utilized in a scheme of social control. Consider the matter of accident prevention. We may seek to eliminate industrial accidents by a program of government regulation

and inspection. Or we may make it financially worth while for employers following their own self-interest to eliminate accidents as far as possible. This the Workmen's Compensation Law does. The lower the accident rate the lower the premiums that have to be paid. This sort of social control, which motivates the employer by appealing to his self-interest, is far more effective than coercion, inspection, and the ordinary methods of law enforcement. We cannot expect that employers following their own self-interest in a regime of *laissez faire* will conduct industry in such a way that desirable social conditions will ensue. But society can set up social arrangements which will induce private enterprise, following its own self-interest, to direct industry toward the ends conceived by society to be desirable. Thus, if society places the financial responsibility for industrial accidents and unemployment upon business men by compelling them to pay the costs of accidents and unemployment insurance, they will find it to their self-interest to buckle down to the task of minimizing these evils. Thus the principle of social control may itself utilize the automatic regulative forces. And even though nonpecuniary motives should sometime, in a less materialistic age than the present, become dominant, self-interest might still be the dominant motivating factor.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "Regulation implies *system* and *order*. A systematic or orderly world is one which is subject to the regulation of definite conditions and forces." Defend or oppose this statement.

2. "That economic activities are regulated is shown by the fact that a fairly definite proportion of our productive power is devoted each year to the production of wheat, shoes, cloth; that the balance of trade between the United States and England remains fairly constant from year to year; that the per cent of the value of manufactured goods paid out in wages is substantially the same year after year." Defend this statement.

The following table gives the percentage ratio of wages to value added by manufacture, in the United States:

YEAR	PER CENT
1899	41.5
1904	41.5
1909	40.2
1914	41.3
1919	42.1
1925	40.1
1929	36.5
1933	36.2

3. "A rate of wages determined by unrestricted individual competition is a 'natural' rate of wages; a rate of wages determined by combination or by legal enactment is an 'artificial' wage." Do you agree? Defend your position.

4. "Thus, strictly speaking, every development of social structure and function from the family to a police force, from the institution of personal property to the provision of public parks and libraries, from the primitive taboo to the most complicated act of Parliament, is alike 'artificial,' that is to say, the product of human intervention, the outcome of human activities. The plain truth is that to apply the antithesis of 'natural' and 'artificial' to social action is sheer nonsense."—BEATRICE POTTER WEBB Do you agree? Why, or why not?

5. "We must remember, again, that if this system [the present economic order] really possesses the vaunted superiority of 'going by itself,' it is certainly not true that it made itself. If it goes by itself it is because the machinery is by this time well established. But when it had to be started—when individual ownership had to be created, with all its attributes of rent and interest and so forth—it required centuries of conquests, a hundred revolutions, a thousand laws, and all the might of kings, nobles, or parliaments. And, to tell the truth, this transformation still goes on without ceasing, so that it would be very hard to discover how much of the so-called natural order remains underneath the existing economic order."—GIDE. Argue for or against this proposition.

6. "Only free competition is right; co-operation [that is, co-operation between competitors] is, *prima facie*, wicked. Co-operation, in the sense indicated, is only a pleasant name for combining to take advantage of your customers."—TAYLOR. Argue for or against this proposition.

CHAPTER IV · The Organization of Production



THE FUNCTION OF THE ENTREPRENEUR

In this chapter we shall consider the function of the entrepreneur and the structural forms of organization through which entrepreneurship is carried on. By entrepreneur we mean the person or group of persons who control the business unit. If the business is a large one, several persons may share jointly in the function of entrepreneurship. At the outset it should be made clear that by control we do not mean management in the *narrower sense* of that term. Very likely the entrepreneurs employ a salaried manager to run the business. Such a manager is simply a paid employee and is in no sense an entrepreneur. He is working under authority and may be dismissed at any time or at least at such time as his contract terminates. The fact that many entrepreneurs manage their own business in no way implies that the details of the management are a part of the entrepreneurial function proper. In many small businesses the entrepreneur not only manages the business himself but also does a great deal of manual labor. Yet no one has ever contended that manual labor is a function of entrepreneurship. A hundred years ago, when most business was small-scale and almost all individual proprietors managed their own business, it was natural that management should be thought of as the essential part of the function of an entrepreneur. In modern society, with large-scale production, division of labor, and high degree of specialization, the management function has split off in large part from the function of entrepreneurship. In large-scale firms the entrepreneurs frequently do not manage the business at all; that function is delegated to a group of salaried officials selected and appointed by the entrepreneurs. Business managers frequently originate policies, but the entrepreneurs retain final control by their power to step in and alter the policy or oust the managers if necessary.

Yet in spite of the fact that it is necessary to distinguish between salaried managers and entrepreneurs, the point must not be overlooked that *control* implies *management*, in the *broader sense* of that term.

The entrepreneur is the final co-ordinator of the factors employed, including even the salaried managers themselves. The going concern is under the final control of the entrepreneur, or group of entrepreneurs, and the ultimate decisions of management therefore are in their hands. In this sense it is quite correct to speak of management as being a function of entrepreneurship.

It is often said that the essential function of entrepreneurship is the assumption of risk. If this statement is intended to mean that entrepreneurs are the only persons who assume the risks of a business, we cannot agree. If the statement is intended merely to mean that the assumption of risk is inseparable from entrepreneurship, of course we must agree. But any human activity entails risks. A wage-earner assumes risks. He assumes the risk of accident and industrial disease. His employer may go into bankruptcy and may fail to pay any wages whatever. The hired manager assumes risks. He runs the risk of failure and loss of reputation, in which case he not only will lose his present position but will find it difficult to obtain another one. Those who lend money to the concern run the risk of possible loss of all or part of the funds advanced. Any person who does anything assumes risks. It is clear, therefore, that risk-taking is not a special characteristic of the function of the entrepreneur.

But does not the entrepreneur at any rate assume a certain special risk? His capital, it is said, is placed in the front-line trench. It withstands the first shock of loss; it serves as a buffer between possible losses and the claims of wage-earners, salaried workers, and bondholders. While this statement is, generally speaking, true, nevertheless it confuses the function of the stockholders with the function of the entrepreneur. To be sure, the entrepreneurs are generally though not necessarily stockholders in the corporate enterprises they direct, but those who control large corporations may own relatively few shares in these concerns.

The entrepreneurs are the persons who are *in control*—who dictate the business policies of an enterprise. In theory this would include, in the case of a corporation, the entire body of stockholders, or at least all the holders of common stock, since all holders of common stock generally have voting power according to the number of shares they hold. But in actual practice the great body of stockholders in large corporations do not look upon their shares of stock as instruments of control at all. They seldom bother to vote their stock, they know little about the affairs of the company, and they exercise no voice whatever

in determining its general policies. In short, they exercise no control whatever over the enterprise. They are in no sense entrepreneurs. The small group of stockholders who dominate the company, elect the directors and officers, and dictate the policies of the enterprise—these persons are the entrepreneurs.

Occasionally even this inner circle of stockholders loses its control. It may be that they have placed themselves in such a position financially that a banking house, to whom the company is under financial obligation, is able to step in and dictate the company's policies. It may be that the banking house can force the establishment of a voting trust which will insure a continuity of such business policies as will serve the interests of the bank until the obligations are met. In this case the bank is the real entrepreneur.

The phrase "dictate business policy" of course does not mean that any group of entrepreneurs have unlimited arbitrary power over any business. The business policies in question are under their control within certain limitations. There are, first of all, the limitations imposed by the laws of the land. Then there are also such forces as custom, the demands of the public, the opposing wills of the organizations, groups, and individuals with whom they deal. Who, indeed, controls American industrial life, for example? It is altogether too simple an answer to say that it is controlled by some ten million entrepreneurs. Millions of these entrepreneurs—the farmers and small storekeepers, for example—have as individuals very little control over the broad currents of economic activity. In a sense they are inconsequential yeomen in a great "commercial feudalism" controlled by powerful financial groups who are the real barons and kings in the modern world. And yet it is certainly naïve to think of modern industry as being completely dominated by a few absolute and all-powerful financial monarchs. Shall we say that the wage-earners, for example, have not any control over industry in modern countries? Certainly they are a power to be reckoned with, and so are the farmers. Control is, in fact, in modern times a diffused thing. It shows itself in the control of property, in the ballot box, in trade-unions, in trade associations, in farmers' organizations, in women's clubs, in public opinion. Control is much broader than entrepreneurship. Entrepreneurship refers to a special kind of control. It operates within a certain range, subject to the broad influence of social control.

FORMS OF ENTREPRENEURIAL ORGANIZATION

We must consider in greater detail the various forms of business organization through which entrepreneurs function. Five will be described briefly: (1) individual proprietorship, (2) partnerships, (3) corporations, (4) co-operative associations, (5) political units, such as a municipality, state, or nation.

Individual proprietorship was the dominant form in the period of small-scale industry, and it is still of great importance wherever the small business unit survives. Agriculture is still nearly as important as manufacturing in the United States, and it is controlled almost exclusively by entrepreneurs who are individual proprietors. There are, all told, approximately six million of them. Retail stores are still operated to a considerable extent by individual entrepreneurs. This is especially true of business in small towns and of corner stores in the residence sections of large cities. In manufacturing there are still some hundred and forty thousand establishments operated by individual proprietors. In the case of stores and farms, especially, the individual proprietor is frequently entrepreneur, manager, capitalist, and manual worker, all in one.

The great bulk of economic activities which we classify under the head of manufacturing, transportation, and finance is carried on by partnerships or corporations, especially the latter. The *partnership* is an association of individuals engaging in a business. It is formed by means of a partnership agreement, which agreement is a contract between the individuals entering into the partnership. Generally, but not always, each member of the partnership furnishes a part of the capital and participates actively in the management of the business. At all events the entire property of each member is liable for all the debts of the firm. This is known as unlimited liability. Each partner assumes risks as great as though he alone were the proprietor of the business. To be sure, the risk is shared jointly by all the partners, and therefore the liabilities of the firm would devolve on any one partner only if all the other partners proved absolutely propertyless. In one respect the risk is greater than that of an individual proprietor, for any partner can enter into a contract relating to the business in question, which is binding upon the firm even though the other members have no knowledge of the transaction and have never given their consent.

Each and every partner may or may not be an entrepreneur. Entrepreneurship entails, let it be remembered, financial control. The silent

partner who has an investment in the business but assumes no active control is not an entrepreneur. Only the partners who have investments in the business and actively control the final policies are, properly speaking, entrepreneurs.

The *business corporation* also is an association of persons, but it is a very special type of business association. The corporation is a creature of the state. A group of persons decide to form a corporation. They draw up articles of incorporation in which they state the name of the corporation, the purposes for which it is formed, the amount and kinds of stock that the company will issue, the location of the main office, the duration of the charter, and the names and addresses of the persons seeking to incorporate. These articles, properly drawn up, are submitted to some state officer, generally the secretary of state; and if they are legally drawn, he approves and files the application, which then becomes the charter of the newly created corporation. The state having granted the charter, the corporation is looked upon as a legal person, separate and distinct from the individual incorporators. Being a person in the eyes of the law, it can sue and be sued, own property, and enter into contracts. The shares of stock are transferable, and therefore the life of the corporation does not depend upon the lives of the individual stockholders. All the stockholders might be killed simultaneously by an earthquake, but the stock would pass to their heirs, and the corporation would live on. It thus has continuity of life separate and distinct from that of the constituent stockholders.

Unlike the partnership, which usually entails unlimited liability upon each individual partner, the individual stockholders who associate themselves together under the corporate form of organization usually escape all personal liability. As a general rule the creditors of the corporation cannot attach the individual property of the stockholders at all. Before the Banking Acts of 1933 and 1935 the stockholders of national banks were liable for the debts of the corporation up to an amount equal to the par value of the stock they held. In some states all stockholders are so liable except the stockholders of corporations engaged in manufacturing and mechanical pursuits. But at most the liability of stockholders is a limited one. The stockholders, of course, always run the risk of losing their investment in the corporation, should it encounter business difficulties and become insolvent.

The stockholders elect the board of directors, who in turn elect the officers and control the policies of the corporation. Usually each stockholder has a vote for every share of stock owned. In theory, therefore,

the corporation is democratically controlled by the entire body of stockholders. In practice, however, the great body of stockholders in large corporations do not exercise their voting power at all. Some large corporations have more than a hundred thousand stockholders. Each of these persons has a right to attend the annual stockholders' meeting, at which the board of directors is elected and general policies decided upon. The uninitiated might suppose that such stockholders' meetings are very large affairs; yet the general rule is that only a few persons attend. These are the people who dominate it. These active dominating persons who actually control the business are clearly the real entrepreneurs.

Stockholders who exercise no control over the business are not, properly speaking, entrepreneurs. Nor are minority stockholders who seek to control the business, but fail to do so, in any sense entrepreneurs. These stockholders are in reality in the position of investors or speculators who have put their money into a business over which they have no control. It is true that they expose their investments to the same risks as do the entrepreneurs proper (that is, the controlling stockholders). But while entrepreneurs necessarily run financial risks, risk-taking alone, as we have learned, divorced from control, does not constitute entrepreneurship.

Sometimes the bondholders become the entrepreneurs. Bondholders are persons who have lent money to the corporation, and the bonds are evidences of the loans made. Ordinarily the bondholders have no control over the affairs of the corporation; but they always have a contingent control. If the interest on the bonds is not paid, or if the principal is not paid when the bonds mature, the bondholders may foreclose on the property of the corporation and assume a position of active control of its policies. Here again certain dominant bondholders are likely to become the real entrepreneurs.

Another form of association is the *co-operative society*. Two types are of chief importance: (1) consumers' co-operation and (2) marketing co-operation.

The first is an association of consumers to carry on a retail-store business. Several local consumers' co-operative societies frequently form a wholesale society which engages in a wholesale business, supplying the goods at wholesale to the various local stores. Such wholesale societies sometimes engage in certain manufacturing businesses, producing a portion of the goods supplied to the affiliated retail stores. Thus the British Wholesale Society manufactures soap, shoes, and

clothing, and owns a fleet of ships in which tea is shipped from their own plantations in Ceylon and wheat from their own wheat fields in Canada.

Co-operative marketing societies are associations of producers interested in marketing their products. The California Fruit Growers' Exchange, the American Cotton Growers' Exchange, potato exchanges, milk producers' associations, and farmers' elevators are examples. Co-operative creameries, while engaged chiefly in manufacturing, are concerned also with marketing.

Co-operative societies differ from corporations in several important respects. Profits are distributed on a patronage basis after a certain fixed dividend of 5 or 6 per cent has been paid on the stock. In the case of consumers' co-operative societies dividends are paid on the purchases made by each member, and in the case of marketing co-operation each producer is paid a dividend based on the volume of his products marketed through the association.

Another important difference between the co-operative society and the corporation is in the matter of control. Co-operative societies are usually much more democratic in this matter than corporations. In the first place, no shareholder in a co-operative society has more than one vote, regardless of the number of shares he owns. In the second place, shareholders in co-operative societies exercise control over the business to a greater extent than do the rank and file of stockholders in a corporation. Entrepreneurship is therefore somewhat more widely diffused in co-operative societies.

Finally, we must consider *political units in the role of entrepreneur*. Municipalities frequently own and operate their street railways, gas plants, electric-light plants, and waterworks. All Western governments operate postal services, which are enormous business enterprises. On the continent of Europe certain of the governments own and operate the railroads, the telephones, and other public utilities. In the United States some states from time to time have undertaken to operate flour mills, elevators, and cement works, and have engaged in hail insurance, farm-mortgage banking, and commercial banking.

Who are the entrepreneurs in these cases? To say that they are the various political units involved is too general. Entrepreneurship is personal, and in any sort of business enterprise it should be possible to discover the persons who are the real entrepreneurs. In democratic political units the voters ultimately control the enterprise. From the immediate standpoint, the political party, or group of parties, in control of the government dominates the situation. The party in turn is respon-

sible to the majority of voters who elected it to office. Yet the minority party and its body of voters usually have large representation in the legislative branch of the government and are by no means without power. This very fact is one of the difficulties in the way of efficient governmental administration of business enterprises. The minority party is likely to do everything in its power to hamper and block the activities of the majority party in order to discredit it. In the corporation the majority absolutely and completely dominate the situation. If the minority do not like the way the dominant element run the enterprise, they can always sell their stock and get out. This situation gives rise to unified and harmonious control. But in the case of the government the defeated minority cannot get out, and they are always present to interfere with the carrying out of a clearly defined, consistent, and homogeneous policy. To be sure, even in private corporations it sometimes happens that there are several groups of powerful stockholders, none of which can obtain a majority control. Compromise is then essential, and a clearly defined consistent policy becomes difficult and perhaps impossible.

But, it may be asked, are not the taxpayers the real entrepreneurs in government-owned enterprises? We have learned that entrepreneurship means financial control and that the assumption of risk is inevitably associated with such control. Although the taxpayers assume the risk, yet the voters have the final control. Here at least it would appear that entrepreneurship does not necessarily entail any assumption of risk.

From one point of view the distinction made is more imaginary than real. In all modern societies all self-supporting voters are taxpayers, though they may not always be aware of it. Every self-supporting individual, though he may not own any property, at least rents a room or a house. As tenant he indirectly pays the tax on the house. Tariff duties, consumption taxes, and so on are finally borne in the main by the general body of consumers, whether property-owners or not. Yet in fact, in so far as modern governments are tending to abandon indirect taxation and to exempt low incomes from taxation, it would seem that control over government enterprises is being divorced from any assumption of risk. On the other hand, it may be argued that every citizen has an equity in the government-owned wealth and the total government income, and both would be affected by the success or failure of government-owned ventures.

The wider the dispersion of ownership and control of a business the more difficult it becomes to locate the entrepreneur. In the individual-

proprietorship business the case is clear and simple. So also in the closed corporation, which is owned by a small group of men or even by a single family. But when we come to the great modern corporations with thousands of stockholders—publicly owned corporations, so-called—the case is very different. And when the state becomes the unit of business organization, the entrepreneur becomes a very elusive person indeed.

Perhaps it may be asked, Is not the concept of entrepreneurship exclusively a private-property concept? Under socialism, with private enterprise eliminated, would not the function of entrepreneurship disappear? It certainly must be admitted that to undertake and control a business in a regime of private enterprise and free competition is quite a different thing from the functions performed by a monopolistic state engaged in an economic activity. It is quite possible that the two functions are so widely different that it is misleading to speak of a completely socialized state as carrying out the functions of an entrepreneur. Obviously the objection to the term is less if the state controls only a few economic activities in an individualistic and competitive society.

CORPORATIONS AND THE CONTROL OF ECONOMIC LIFE

Many types of production are still carried on under the individual entrepreneur or partnership form of business organization. A farm is in the majority of cases an individual or family enterprise. Retailing, particularly of foods, is usually an individual enterprise or it is carried on by small corporations the stock of which is owned in each case by a few persons. But large-scale industry is dominated by the large corporation. All railways and all public-utility concerns are corporations and so too are most large industrial concerns. The giants among corporations are found in all three types of concern: the Pennsylvania and the New York Central railway systems, the American Telephone and Telegraph Company, and the United States Steel and General Motors corporations.

The giant corporation has come to occupy a position of very great importance in modern business. It dominates transportation, the electric light and power industry and the so-called large-scale industries such as iron and steel, motors, electrical equipment, and machinery. It has made some headway in the field of merchandising. Berle and Means found that in 1930 the 200 largest corporations (other than banking concerns) controlled between 15 and 25 per cent of the national wealth

and between 35 and 45 per cent of the business wealth of the United States. The table below shows the control of these corporations.

From this table it can readily be seen that the most important form of control in 1929 was management control. It is difficult to determine precisely how the managers—that is, the superior officers—obtain and perpetuate their power over the affairs of the corporation. The mere mechanics of the operation, however, are clear. Since the directors and officers do not control the stock through ownership, they must control enough votes of stockholders to dominate the annual stockholders' meeting, at which only a very small percentage of the outstanding stock is customarily voted. This control is obtained by asking the individual holders to assign their voting privileges to a proxy committee, which then votes the stock at the annual meeting. Now the proxy committee is named by the management and therefore votes the stock as the management demands. How are the proxies of the individuals obtained? Presumably simply by request from the management. Persuasion, personal relations, and other methods are employed.

TABLE 7. Percentage Distribution of the Types of Control among the 200 Largest Nonbanking Corporations in the United States, 1929¹

TYPE OF CONTROL	PERCENTAGE BY NUMBER	DISTRIBUTION BY WEALTH
Management control	44	58
Legal device	21	22
Minority control	23	14
Majority ownership	5	2
Private ownership	6	4
In hands of receiver	1	negligible

There are several legal devices by which a minority group can obtain control of a corporation. The holding company is one such device. For example, in such a company the first holding company acquires 51 per cent of the common stocks of one or more operating companies, then another holding company holds a controlling amount of the stock of this and other similar companies, and so on until sometimes the control company is removed five or more times from the operating company. Now any group of men that can control the topmost company can control the operating companies from which all the earnings come. It is evident that this process of diluting the ownership makes it possible

¹From A. A. Berle and G. C. Means, *The Modern Corporation and Private Property*. By permission of the authors and the publisher, The Macmillan Company.

for those who have actually invested a very small percentage of the total common-stock capitalization of the operating companies to control them. This device has been widely employed in the field of public utilities.

Another legal device for separating the investor from the control of the company is the issue of different classes of shares. The promoters of Dodge Brothers, Inc., bought the property for a sum reported to have been about \$150,000,000, and issued \$160,000,000 of bonds and preferred stock against it. They then put out 1,500,000 shares of Class A stock and 500,000 shares of Class B stock. But only the latter carried voting privileges. It is not correct to say that the owners of the Class A common stock were deprived of control. They bought into this fancy arrangement with their eyes open. By the use of this same device it was possible for the Van Sweringens to control a railroad with less than one half of one per cent of the total investment in it—virtually no interest. The New York Stock Exchange has now refused to list common stock thus made impotent of voting power.

A third legal contrivance for separating the investors from control is the voting trust. Here the trust, which is usually formed by the management, acquires some stock and offers to the public the trustee certificates issued by the trust which entitles the holder to whatever dividends the trustees choose to distribute. With the funds derived from the sale of the certificates more stock can be bought, more certificates issued and sold, and the process continued until the control group in the trust have acquired enough voting rights to dominate. Such voting trusts are usually of limited life, but when one expires another may be organized. This device is not very widely used because it places open responsibility upon the trustees and subjects them to certain legal liabilities that are not placed upon the control group by the other devices.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The entrepreneur is therefore the point on which the whole of the economic mechanism turns. On him converge all the factors of production; and from him also, as we shall see, proceeds all the income. . . . The entrepreneur, therefore, is both the great employer and the great distributor."—GIDE. Comment.

2. "They [entrepreneurs] represent that class of people who make rather than take a job. From one standpoint, these are the only two important types of workers anywhere in the world. We have men who attend to a task after it is pointed out to them and others who create their own task and continue developing it as long as they live. Wage-earners, generally speaking, though

with notable exceptions, belong to the first group, and entrepreneurs to the second.”—BOUCKE. Is this distinction sound? Does it properly define “entrepreneur”?

3. “He [the entrepreneur] has been well called a captain of industry, for he commands the industrial forces, and upon him more than any one else rests the responsibility of success or failure.”—ELY. Do you agree?

4. “In the last analysis, then, the majority stockholders, who control the policy of the company, are the chief enterprisers in the case; although, sharing the risk and responsibility with other stockholders and officials, they share the profits also with them.”—LE ROSSIGNOL. Argue for or against this statement.

5. Do you think it would be better to say merely that entrepreneurship is one of many forms of human activity of economic significance, than to set it up as a fourth or special factor of production?

6. As usually managed, the large corporation represents feudalism in business. Comment.

7. Why might a small investor prefer to own the stock rather than the bonds of a corporation the policies of which he had no power to control?

CHAPTER V · Combination of the Agents of Production



THE BACKGROUND OF SOCIAL INSTITUTIONS

We have noted in a previous chapter that all physical goods and all services consumed by man come ultimately from two sources: from man himself and from the environment external to man. There are, then, only two primary means of production, man and nature. But the extent to which man can utilize the resources of the region in which he lives is affected by his power to command these natural forces. It is a commonplace to say that the high level of well-being of the people of the United States at the present time is owing to the richness of the land that they inhabit. Yet the American Indian had exactly the same natural resources for his environment as we have and still lived at a low standard of comfort. But he made little use of the resources in comparison with the extent to which we have utilized them. In part this difference is doubtless to be explained by the higher innate capacity of the white race for the type of civilization that prevails in the Western world. But to an even greater extent it is owing to the fact that we possess a body of knowledge, derived from scientific discovery and invention, which enables us to command and control the resources latent in our environment. Also in Europe, where essentially the same racial groups now possess the land as in earlier centuries, production is much greater than it was two hundred years ago. Life is easier, and people have, on the average, many more goods to consume.

There is thus a third element or condition, in addition to nature and man, governing the productivity of a people. This condition is customarily referred to as "the state of the arts." But we must be careful not to think of the state of the arts as a third primary means of production. It is a powerful force in determining the productive capacity of man and nature, but not a concrete, tangible, objective thing like man and nature. The state of the arts is the medium in which men operate; it is not material from which they produce. Stated less abstractly, the state of the arts determines the extent to which men are able to draw forth and turn to their own purposes the resources of the earth and of mankind.

Still another condition must be mentioned as an element in determining the productive capacity of man and nature. The nature of the social organization of a people, including the forms of government, legal systems, attitudes toward government, and much besides, also determines whether they will get much or little out of their environment. Earlier civilizations differed from our own quite as much in the extent to which men were able to co-operate as they did in the extent of scientific knowledge. They differed from our own in the ways in which individual initiative was stimulated or retarded. They differed also in the main objectives of human activity. But these conditions are no more means of production than is the state of the arts. Like the latter they affect productivity powerfully, but they are not concrete physical agents that can be used for making goods. They are qualifying circumstances that vary as time passes, and they act through the modification of man's behavior.

THE FACTORS OF PRODUCTION

Production, as we have noted, is carried on by man and nature. Human beings perform three types of service, each of a different order. These three types of service are (1) labor, (2) entrepreneurship, and (3) saving. By labor we mean any human activity, either of body or of mind, that leads to or assists in the production of economic goods or the rendering of economic services. Entrepreneurship involves the assumption of the responsibility for carrying on and controlling individual business enterprise and, indirectly, all productive activity in society. Saving makes possible the indirect, or roundabout, method of production. By this method capital equipment is made first, and this in turn is used to produce consumers' goods. Saving enables labor and entrepreneurship working with natural resources to produce capital goods.

In the study of production the term "capital" is employed in a special sense. It includes all man-made agents of production. It does not include natural resources; nor mere evidences of property rights, such as stocks and bonds. Capital consists of concrete material goods and not of legal claims. Consumers' capital consists of finished goods ready for consumption, such as houses, musical instruments, libraries, and furniture in the hands of consumers. Producers' capital consists of wealth that is still in process, such as raw material, or of aids in production, such as machinery, factory buildings, and the equipment of railroads. But the major part of all capital is used by producers. Hence

we shall generally use the term to mean producers' capital, and when consumers' capital is meant that meaning will be clearly indicated.

The concrete items that compose producers' capital are not exactly alike. Some, like factory buildings, are very *durable*, and others, like coal, are more *perishable*; some, like iron ore, are *unspecialized* and can be used to aid in the production of a variety of goods, whereas others, such as spinning machines and blast furnaces, can be used at only one step in the production of one commodity or a closely related group of commodities.

We find then that there are four essentially different types of service-rendering objects which may be distinguished from the point of view of economics: labor, natural resources, capital, and entrepreneur. In the present state of the arts all these factors are necessary for production, although it is possible to imagine a society so primitive that capital and entrepreneurs would be entirely absent.

Since all are necessary to carry on the modern productive process, it follows that productivity will be affected very greatly not only by the sum total of them but also by the relative amounts of each that are available for combination with all the others. In the evolution of industrial society many different combinations have been made. In sparsely settled regions much land has ordinarily been combined or used with little labor, labor being the scarce factor. Before the Industrial Revolution changed so profoundly the state of the arts, capital was of little importance in the combination of the factors. And with the development of large plants and giant corporations the number of independent entrepreneurs relative to the total population has decreased. In farming alone, the independent entrepreneur continues to hold his own.

THE RELATIVE SCARCITY OF THE FACTORS

The way in which the four factors of production are combined depends in a large measure upon the supply of each. For example, whatever capital equipment is in existence will certainly be utilized somehow in the long run unless it has become entirely obsolete; and the manner in which it will be combined with the other factors will depend a great deal upon its relative abundance or scarcity compared with the other factors.

Of the four factors of production, natural resources are limited in quantity by nature. Population is still growing, and from it springs an increasing supply of entrepreneurs and of laborers both manual and

mental. Capital is rapidly accumulating. It is true that during the nineteenth century the land supply also was increasing. The railroads opened up vast areas of land in North and South America, Australasia, Russia, Africa, and parts of Asia. This phenomenal exploitation of virgin natural resources is one of the most important characteristics of the last hundred years.

But certain individual countries long ago experienced a limitation in the supply of the natural factor in production. Thus England, at the close of the eighteenth century and during the first quarter of the nineteenth century, was experiencing grave difficulties because of the fact that while her population was increasing, her food-producing areas were already limited. Because of these difficulties there developed a considerable public discussion that resulted in the gradual formulation of the principle known in economics as the law of diminishing returns.

THE LAW OF DIMINISHING RETURNS

The following is a statement of this law as given by John Stuart Mill, the great English economist, writing in 1848. "After a certain, and not very advanced, stage in the progress of agriculture, it is the law of production from land, that in any given state of agricultural skill and knowledge, by increasing the labour, the produce is not increased in an equal degree; doubling the labour does not double the produce; or to express the same thing in other words, every increase of produce is obtained by a more than proportional increase in the application of labour to the land."

This is a statement of the law of diminishing returns applied to land. In fact, from a social point of view, the application of the law to land or natural resources is of primary importance.

Let us illustrate this principle concretely. Suppose we take 300 acres of land, which we will consider the fixed, or constant, factor. Suppose there is applied by way of experimentation an increasing quantity of variable factors to this plot of land. The results of this hypothetical series of experiments are shown in columns 3 and 4 of Table 8.

It is evident from Table 8 that in this series of imaginary experiments the point of diminishing returns is reached when the labor of three men with necessary equipment is applied to 300 acres of land. From this it would follow that even if land were so abundant that it was a free good, men would not cultivate more than 100 acres per laborer; for if the work were spread out thinly over more land, the

TABLE 8. Combination of Factors of Production

1	2	3	4	5
FIXED FACTOR: 300 ACRES OF LAND	VARIABLE FACTORS: LABOR AND CAPITAL	PRODUCT TOTAL	AVERAGE PRODUCT PER UNIT OF VARIABLE FACTORS	MARGINAL PROD- UCT: ADDITION TO PRODUCT DUE TO LAST ADDED INCRE- MENT OF VARIABLE FACTORS
Acres 300	1 man with normal equipment	Bushels 3,000	Bushels 3,000	Bushels
300	2 men with normal equipment	7,000	3,500	4,000
300	3 men with normal equipment	10,800	3,600	3,800
300	4 men with normal equipment	14,000	3,500	3,200
300	5 men with normal equipment	16,000	3,200	2,000

returns would yield a smaller product per laborer. If 300 acres per laborer were cultivated, the product would be only 3000 bushels; at 150 acres per laborer the product would be 3500 bushels; at 100 acres, 3600 bushels. So far intensive cultivation is more profitable than the extensive cultivation of a plot too large to permit proper attention and care. If, however, only 75 acres per laborer were used, the per capita product would decline to 3500 bushels. If 60 acres per laborer were utilized, the per capita product would be only 3200 bushels.

If, therefore, there were an abundance of free land, all of equal grade, not more than 100 acres per laborer would be used, no matter how small the population of the country. But suppose the population is so large that there are not 100 acres per laborer. It may be that there are only 75 acres per laborer. Under this situation it is obvious that the natural resources will have to be utilized beyond the point of diminishing average returns.

In frontier communities no doubt it frequently has happened that labor and capital were spread out over so much land that the land was not utilized even to the point of diminishing returns; or to put it an-

other way, they were not utilized up to the point of highest per capita returns. This sort of cultivation was certainly uneconomical, but persons familiar with frontier agriculture can testify that such wasteful and ill-advised methods were used. As population increased, the point of diminishing average returns was reached. As population became denser, cultivation became more intensive, more labor and capital were used per acre, and agriculture was pushed far beyond the point of diminishing returns.'

None of the other factors are limited in the manner in which certain natural resources are limited. The amount of arable land is more or less fixed, though there is great variation in the fertility of the available supply. The same is true of water power, of iron, coal and other mineral deposits, and of timberland.

The numbers of entrepreneurs and scientific experts are limited partly by biological heredity, partly by educational and industrial opportunities. Only a certain proportion of each generation is fit timber for entrepreneurship, for the higher grades of managerial labor, and for the highest type of scientific work. There is always more or less movement from the labor group to the entrepreneurial group or vice versa. It must be remembered that labor as we here use the term includes the highest-grade salaried worker as well as the more ordinary clerical worker and manual laborer. The margin of entrepreneurship, that is, the point of indifference as to whether one is a salaried worker or an entrepreneur, is a continually shifting one. Superior entrepreneurs tend to bid natural resources, labor, and capital away from marginal entrepreneurs (entrepreneurs who are now making no profits above wages and interest on invested capital), thus rendering it impossible for them to remain entrepreneurs. The less efficient entrepreneurs are squeezed out and pushed into the ranks of wage-earners or salaried workers. Thus a new margin of entrepreneurship is established.

The supply of entrepreneurs relative to the supply of laborers therefore is flexible. There is a constant ebb and flow between the wage and salaried group and the entrepreneurial group. If superior entrepreneurs are unable to utilize all the available factors of production, inferior entrepreneurs will come into the field of entrepreneurship from the ranks of the wage-earners or salaried workers. Thus the void can readily be filled. Accordingly the supply of entrepreneurs is as expansible as is the supply of labor. But with the continued slowing down in the rate of population growth the labor supply is now increasing much less rapidly than it was in the nineteenth century.

With respect to the other factor of production, capital, it is astonishing how rapidly it accumulates in modern society. The last few centuries, with newly discovered undeveloped resources, improvements, and inventions, have opened up a field for capital investment so vast as to preclude seemingly any possibility of sufficiency of capital accumulation, and yet sufficient capital has been forthcoming year after year. In fact, the increased production consequent upon these discoveries and inventions is itself the source, in large part, of this capital accumulation.

As we proceed with our study we shall see the important bearing of the principle of diminishing returns from natural resources upon economic rent, wages, land values, value of farm products, population, immigration, taxation, and upon a host of other problems.

AGENTS OF PRODUCTION

The productive process has been discussed thus far from a rather broad viewpoint. We have considered the factors of production in general terms, as labor, natural resources, capital, and entrepreneur. And the combination of these factors has been treated from the social point of view.

In the actual process of production, as we find it in individual businesses and as it appears to the entrepreneur who directs a business, the factors of production are not so simply classified. The entrepreneur thinks in terms of common labor, machine operators, foremen, clerks, departmental managers, and the like, and not in terms of labor in general. When he considers a combination of productive forces, he often considers the economic advisability of adding a kind or grade of labor to other kinds (or subtracting), as well as adding labor in general to capital and natural resources. Likewise he breaks up the general factor—capital—into many subclasses and combines power units with machine units, or both with raw materials. In the same way natural resources are classified and combined, not as a factor, but as several different varieties.

When the factors of production are thus split up into various subclasses, we shall refer to them as *agents of production*. On page 61 is a schematic representation of the means of production and their relationship to the four factors, and a few of the agents into which they are classified in the individual business.

THE BEST COMBINATION OF THE AGENTS OF PRODUCTION

Under this head we shall consider the problem of production as it confronts the entrepreneur in the individual business establishment. This problem consists of two parts. First, in what proportions shall he employ the various types of labor, natural resources, and capital goods that are available in the market? Here the problem is to find the most profitable combination of the employed agents. Secondly, how large a quantity of these agents shall he employ in conjunction with himself as entrepreneur? This has to do with the problems of large-scale production and the size of the business unit. Taken together, the two problems are concerned with the most profitable combination of all the agents of production, including entrepreneurship itself.

COMBINATION OF AGENTS OF PRODUCTION

PRIMARY MEANS	FACTORS OF PRODUCTION	AGENTS OF PRODUCTION
Man	Entrepreneur . . .	Varying sorts of entrepreneur
	Labor	Common workers Machine operators Semiskilled workers Skilled workers Clerical workers Foremen Executives
Nature	Capital goods . . .	Buildings Other structures Prime movers (engines, motors) Machines Means of transport Raw materials Finished goods awaiting consumption
	Natural resources .	Building land Roadways, waterways Mines Forests Fisheries Crop land Pasture land Recreation sites

An entrepreneur finds that each of the various agents he may wish to employ has a certain market value. In a later chapter we shall see how these valuations are established, but for the present we shall accept

them as established, without inquiring into the reason why. An entrepreneur finds that he must pay established standard wage rates for the various grades of labor that he wishes to employ; he must pay certain prices for capital equipment and raw materials; and if he wishes to rent or to buy natural resources, he must pay the going rentals or prices.

The problem of the best combination of the agents of production must be distinguished from the social law of diminishing returns. The latter deals with great groups of agents, and it assumes the physical supplies of the factors of production without reference to their cost. It simply says, If you have a limited land supply, and if your population and capital increase, you will find that the average physical product per unit of these variable factors will increase at a less than proportionate rate unless the additions to population and capital happen to be accompanied by a change in the arts that makes available more natural resources or increases the efficiency of human labor.

The problem of proportionality of agents refers to the internal management of a plant. It assumes that the prices of the agents have been set by the market outside the plant. It asks, How should these agents be combined to get the maximum money returns from them, present prices assumed? It is a narrower and somewhat different application of the wider and more general law of diminishing returns.

If we assume as fixed the various prices that must be paid for the agents of production, how many of each will the entrepreneur employ? Let us take a concrete illustration. Suppose that an entrepreneur has built a shoe factory of a certain size. The factory building and the land on which it stands we shall call the fixed, or constant, factor. To be sure, an adjoining piece of land might be purchased and an addition built to the factory building; but except for such additions, which would be made only at long intervals of time, if at all, the factory building and the land on which it stands form a constant, or fixed, factor. To this fixed factor variable factors may be added. Shoe machinery may be installed, labor may be employed, and raw materials purchased. At first the entrepreneur may install a relatively small number of machines and employ relatively little labor. Gradually, as he builds up a wider and wider clientele of customers, he will install more machinery and employ more labor. He will apply more and more variable agents to his fixed plant. What will be the effect of these successive combinations of agents upon his cost per unit of output?

AVERAGE PRODUCT AND MARGINAL PRODUCT

Tables 9 and 10 illustrate the principles involved in this problem. It should be noted that in these tables rather large jumps are made from any one given situation to the next. Ten units of variable agents are successively added. These units of variation are fairly large, and accordingly the results obtained, in terms of product and cost, give a set of discontinuous figures. For this reason it is impossible from these rough tables to determine precisely the *point* of lowest total unit cost. The smaller the increments of variation the more nearly would the marginal unit cost equal the total unit cost at its lowest point. As will be seen, they do not precisely correspond in the tables because of the unavoidable discontinuity involved in the tabular method of presentation. Later in the chapter these relationships will be presented in the form of curves. By the use of the diagrammatic method infinitesimally small increments of variation are represented and thereby precise relationships established. The value of the arithmetical tables here used is that they enable the student, by means of simple calculations, to grasp the fundamental principles involved. Table 9 gives the different numbers of agents employed in each successive experiment with the resulting *total product*, *average product per increment of variable agents*, and *marginal product*, that is, the product added by the last increment of the variable agents. The problem is simplified by lumping all the variable agents together as though they were homogeneous. In reality, of course, the variable agents consist of a complex assortment of different kinds of machines, tools, and raw materials, and scores of different kinds of specialized labor.

From Table 9 it appears that at first as variable agents are added the total product increases more rapidly than the increase in the variable

TABLE 9

FIXED AGENTS	VARIABLE AGENTS	TOTAL PRODUCT	AVERAGE PRODUCT PER 10 VARIABLE AGENTS	MARGINAL PRODUCT PER 10 VARIABLE AGENTS
Factory building, land, and entrepreneur	10	1000	1000	
	20	2200	1100	1200 ✓
	30	3375	1125	1175 ✓
	40	4440	1110	1065
	50	5420	1084	980
	60	6250	1042	830
	70	6800	971	550

agent. In other words, the point of diminishing returns has not yet been reached. This point will be reached, however, when 30 units of variable agents are employed. From this point on, the increase in the total product is less than proportional to the increase in the variable agents. These points are brought out clearly in the average-product column. The last column gives the marginal product; that is to say, the product added by the last increment of variable agents. The second 10 units of variable agents added 1200 units of product. The last 10 units of variable agents employed added 550 units of product to the total.

COST PER UNIT OF OUTPUT

The individual entrepreneur must pay the going rates of wages, interest, and rent. He must likewise pay the established market prices for materials and equipment. But the cost of the *finished product* is not definitely fixed. A variation in the combination of the agents will result in a variation in the amounts of finished goods produced and so alter the cost per unit of output.

How far will it pay the entrepreneur to continue to add variable agents to his fixed plant? It is impossible to answer this question without knowing what he will have to pay for each of the agents he employs. Suppose he pays a fixed rent of \$500 a month for the plant; suppose he pays himself \$500 a month as entrepreneur and general manager; and suppose he pays \$100 a month for each unit of the variable factors. We can then determine the cost of each unit of output. These results are given in Table 10.

TABLE 10. Cost per Unit of Output

VARIABLE AGENTS	OUTPUT	FIXED PLANT AND ENTREPRENEUR	VARIABLE AGENTS	TOTAL COST PER UNIT OF OUTPUT
10	1000	\$1.00	\$1.00	\$2.00
20	2200	0.45	0.91	1.36
30	3375	0.30	0.89	1.19
40	4440	0.23	0.90	1.13
50	5420	0.18	0.92	1.10
60	6250	0.16	0.96	1.12
70	6800	0.15	1.03	1.18

In a successful business the money derived from the gross sales must cover at least interest and depreciation charges on the fixed plant, as well as wages, cost of raw materials, interest on investments and

machinery, and many other expenses arising from the use of the variable agents. The output must sell for at least enough to cover the expense of the fixed plant, the services of the entrepreneur, and the variable agents. With every increase in output the cost of the *fixed agents* becomes less and less per unit of output. On the other hand, the cost of the *variable agents* per unit of output will necessarily decrease up to the point of diminishing average returns and will increase beyond that point. So long as the product increases more, proportionally, than the increase in the variable agents, the variable cost per unit of output declines; and when the product increases less, proportionally, than the increase in the variable agents, the variable cost per unit of output will obviously rise.

Now the entrepreneur will not cease adding variable agents at the point of diminishing average returns. In our illustration he will, in fact, if he is thoroughly enlightened, continue adding variable agents until he employs 50 units of them. The reason for this becomes clear when we observe that as the entrepreneur increases his use of variable agents from 30 units to 50 units the increased cost of the *variable agents* per unit of output is more than offset by the decreased cost of the *fixed agents* per unit of output. This will be seen from an examination of the total-cost column in Table 10. The total cost per unit of output is the lowest when 50 units of variable agents are employed and the output is 5420. Beyond this point the increase in the cost of the variable agents per unit of output is greater than the decrease in the cost of the fixed agents per unit of output, and the total cost rises. The entrepreneur naturally will wish to produce at the lowest possible unit cost. He therefore, if he follows a perfectly enlightened policy, will employ 50 units of variable agents.

LOWEST UNIT COST FOR MARGINAL ENTREPRENEURS

According to Table 10 the entrepreneur can produce 5420 units of output at a cost of \$1.10 per unit. This is the lowest total unit cost at which he can operate, and in the long run the price at which the output sells in the market must be at least as high as that or he will be forced out of business. At that price he just covers all the costs he has incurred and receives for his own work the same return he could reasonably expect if he were employed as a manager by some other firm. When an entrepreneur is just able to cover his expenses in the manner described, he is called a marginal entrepreneur, and the firm which he manages is said to be a marginal firm.

Now we have shown that the marginal entrepreneur will employ more and more variable agents until he finally reaches the point of lowest cost per unit of output. Or to put it in another way, he will continue to employ more and more variable agents as long as they more than pay for themselves. If we revert to Table 9,¹ we shall notice that while the second 10 units of variable agents cost \$1000, they added 1200 units to the product. At \$1.10 per unit this will amount to a value product of \$1320. Obviously these second 10 units of variable agents more than pay for themselves. The third 10 units of variable agents also cost \$1000, and they added 1175 units or a value of \$1293. The fourth 10 units of variable agents added a value of \$1172; the fifth, \$1078; and the sixth, \$913. It is clear that our entrepreneur could not afford to employ the sixth 10 units, because they added a value less than their cost. If our entrepreneur had added the sixth 10 units of variable agents, he would thereby have raised his total unit cost above the price of \$1.10. On this basis he could not afford to remain in the business. He would be driven out by those entrepreneurs who were wise enough to employ the variable agents only up to the point where they paid for themselves, or, in other words, the point of lowest total unit cost.

LOWEST UNIT COST FOR SUPERIOR ENTREPRENEURS

There are, however, in every line of business some entrepreneurs who are able for various reasons to turn out goods at total unit costs more or less below the corresponding costs of the marginal entrepreneur. This superiority with respect to unit costs may be the result of a combination of favorable circumstances, frequently of an accidental or temporary character. The lower cost might be due to the invention of a new machine which is covered by a patent; to the introduction of improved methods of routing the materials through the plant, or to a new system of wage payments resulting in higher worker productivity; it might be due to the lucky purchase of an advantageous site at an exceptionally favorable price; or to the monopoly control of a richer source of supply of raw materials. But these are at best only temporary causes of lower unit costs. Innovations will be adopted by others. The profits from advantageous sites or from the control of exceptional sources of supply will be capitalized; a new buyer, for example, of such sites or sources of supply will be compelled to count the

¹The student can readily make this calculation by multiplying the units given in the last column by the price per unit, namely, \$1.10.

price paid for them as a part of his cost of production. All these illustrations of lower cost therefore tend to disappear under the pressure of competitive and price-making forces.

In a later chapter we shall consider differential costs between various competing firms in just such disequilibrium situations. Differences in unit costs are due to the failure of instantaneous adjustments to the changed conditions continually arising in a highly dynamic society. Here we are not concerned with the dynamics of this problem. At this point we are interested only in the fundamental principle of the effect of various combinations of agents of production upon unit cost.

We note the fact that in a dynamic society the low-cost entrepreneur is able to get a larger total output from the employment of a given quantity of agents of production than the marginal entrepreneur. In other words, the low-cost entrepreneur is able by reason of his innovations to make a better utilization of the agents than is made in the plant of the marginal entrepreneur.

Will the superior entrepreneur use the same combination of agents as is used by the marginal entrepreneur? Will he employ the variable agents farther beyond the point of diminishing returns than the marginal entrepreneur does? Again let us take a concrete illustration. We must remember that the superior entrepreneur is able to get a larger output out of equivalent quantities of agents than is the marginal entrepreneur. Let us assume that he gets 10 per cent more. Since he gets a larger product from the agents employed than the marginal entrepreneur, it follows that his total unit cost will be less for each application of the variable agents. He therefore can afford to employ the variable agents beyond the point of lowest total unit cost, since he is able to sell the product at the same price as the marginal entrepreneur. Like the marginal entrepreneur, he will continue to employ the variable agents as long as they at least pay for themselves. Table 11 shows the value of product added by each successive amount of variable agents. It will be remembered that the product is 10 per cent higher than for marginal entrepreneurs.

It will be seen from Table 11 that the superior entrepreneur can afford to employ a total of 60 units of variable agents. Our marginal entrepreneur could afford to employ only 50 units of variable agents. The marginal entrepreneur could get only 830 units out of his sixth increment of variable agents; and since this yielded a value product of only \$913 while the cost of the added increment of agents was \$1000, it would not pay him to employ the sixth increment of variable agents.

But the superior entrepreneur could get 913 additional units out of the sixth increment of variable agents, and this would yield a value product greater than the cost. The superior entrepreneur therefore may be expected to employ the variable agents *beyond the point of lowest total unit cost*, since the largest volume of profits can be made only by extending production in the various plants until the marginal cost equals the selling price. At this point the total unit cost will be somewhat raised, but the higher cost will be more than offset by the larger volume of sales. So long as the marginal cost is below the selling price, every addition to output will add something to the total volume of profits. The net profits are largest at the point where the marginal cost is equal to the selling price.¹

TABLE II

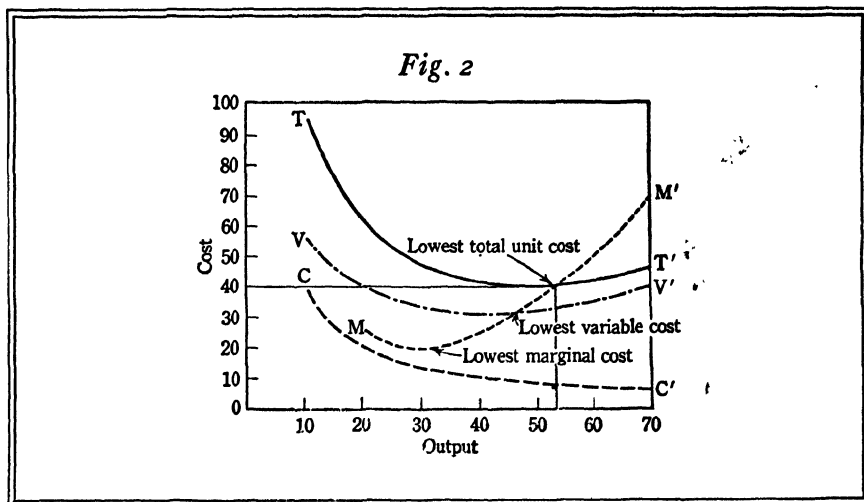
VARIABLE AGENTS	TOTAL PRODUCT	PHYSICAL PRODUCT ADDED BY LAST INCREMENT OF VARIABLE AGENTS	VALUE PRODUCT ADDED BY LAST INCREMENT OF VARIABLE AGENTS	COST OF ADDED VARIABLE AGENTS	NET GAIN FROM EMPLOYMENT OF ADDED VARIABLE AGENTS	TOTAL PROFIT (+) OR LOSS (-)
10 . . .	1100					\$- 790
20 . . .	2420	1320	\$1452 ²	\$1000	\$452	- 338
30 . . .	3713	1293	1422	1000	422	+ 84
40 . . .	4884	1171	1288	1000	288	+ 372
50 . . .	5962	1078	1186	1000	186	+ 558
60 . . .	6875	913	1004	1000	4	+ 563
70 . . .	7480	605	666	1000	- 334	+ 228

The principles stated above can be illustrated more clearly by the aid of diagrams. The tables developed our principles by showing how the output is affected by large changes in the quantity of variable agents used. We jumped from 10 units of variable agents to 20 units, from 20 to 30, and so on. The tables, as we have seen, consist of discontinuous series. In the diagrams, however, the series are continuous, and we can therefore determine the precise relationship at each output point of constant, variable, marginal, and total unit cost.³

¹At this point the student should calculate the marginal unit cost. This he can do by dividing the cost of the added variable agents (5th column in Table II) by the physical product added by the last increment of variable agents (3d column in Table II). This calculation will show that the marginal unit cost is \$1.095 at the point at which 60 variable agents are employed. If infinitesimally small increments of variation had been used, the marginal unit cost would be exactly equal to the selling price.

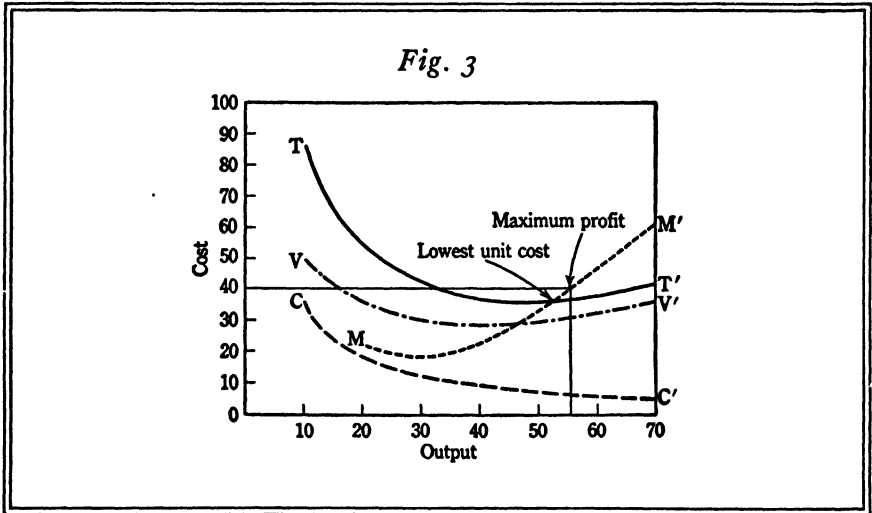
²The price per unit is \$1.10.

³See Alvin H. Hansen, "Prime Costs in the Business Cycle," *Journal of Political Economy*, February, 1924.



In Fig. 2 four cost curves are represented. These curves are, of course, used to illustrate a general principle, and the configuration of similar curves would vary greatly with different plants and industries. CC' represents the cost per unit of output in terms of the constant agents. The larger the output, the lower is the share of the constant cost attributable to any one unit of product. VV' represents the cost per unit of output in terms of the variable agents. It will be noted that this curve at first falls until the point of diminishing average returns is reached, after which it rises. TT' represents the total unit cost. It is derived by adding the per unit constant cost to the per unit variable cost for any given quantity of output. This can readily be seen by inspecting the diagram. At the output of 40 units, for example, the per unit constant cost is 10, while the per unit variable cost is 31. The total unit cost is therefore 41. Similarly, at the output of 53 units, the per unit constant cost is 8, and the per unit variable cost is 32, making the total unit cost 40. MM' represents the marginal cost, or the added cost incurred in producing the last increment of product. An examination of the diagram will show that MM' intersects VV' at the point of lowest average cost of the variable agents—that is, at the point of diminishing average returns. This necessarily follows from the fact that so long as the marginal cost is below the average unit cost in terms of the variable agents, any addition to the output reduces the per unit variable cost. On the other hand, as soon as the marginal cost reaches the average variable cost, any further increase in output resulting from additional applications of the variable agents will raise the average

variable costs. Therefore MM' necessarily intersects VV' at its lowest point. It will also be noticed that MM' intersects TT' at its lowest point. This necessarily follows from the fact that so long as the marginal cost is below the total unit cost, each addition to output lowers the total unit cost, whereas if the marginal cost exceeds the total unit cost, every addition to output increases the total unit cost.



As we have already shown, marginal firms may be expected to expand production under the pressure of competition to the point of lowest total unit cost, but no farther. This lowest total unit cost for marginal firms tends by competition to become the prevailing price. Superior firms will be able to produce at a lower total unit cost. But superior firms will tend to expand production beyond the point of lowest total unit cost. As was shown in Table 11, they can profitably expand production until the marginal cost equals the selling price. The result of this expanded production will be to raise the total unit cost somewhat, but the higher cost will be more than offset by the larger volume of business. In other words, it is clear that so long as the marginal cost is below the selling price, every addition to output will add something to the total volume of profits. Beyond this point it would be unprofitable to go, since the cost of the additional output would exceed its value. These points are illustrated in Fig. 3.

There will be a tendency, therefore, for both marginal and superior firms to continue to add variable agents until the marginal cost equals the selling price. For the marginal entrepreneurs this will coincide with

the point of lowest total unit cost. For superior entrepreneurs this point will be beyond the lowest total unit cost. However, the total unit cost for superior entrepreneurs at the point of maximum profit will be more or less below the selling price, depending on the degree of superiority.

Our superior entrepreneur may be able to utilize effectively a larger fixed plant (either in the form of a larger factory or more factories) than the marginal entrepreneur can successfully control. In other words, he will not only utilize his fixed plant more intensively (by the employment of more variable agents), but he will also tend to utilize his own entrepreneurship more intensively. But entrepreneurial capacity is itself a limited factor; and as more and more plants are added, the point will at last be reached at which it will not be economical to take on any more plants. This leads us to the subject of large-scale production and large-scale management.

LARGE-SCALE PRODUCTION

We are now prepared to consider more fully the principles that determine the size of an establishment, to answer the question why large-scale production prevails in some industries while in other industries the small-scale plant prevails. We must distinguish sharply between two things that are likely to be confused: (1) the size of the firm, and (2) the size of the plant. We may have two types of large firms: one, the firm owning many small establishments, such as the chain stores or the chain restaurants; the second, the firm owning one or more large-scale plants or establishments. There are two tests for determining whether a plant is operating on a large scale. It may be a large establishment because of the considerable amount of capital invested, though the number of wage-earners may be small. An example of this type is the modern giant flour mill, employing expensive equipment but relatively few wage-earners. On the other hand, if many wage-earners are employed, the establishment will be considered a large one, even though relatively little capital is invested. An example may be found in the manufacture of cheap jewelry. Occasionally there are small plants, such as butter factories, that have no large total investment of capital but nevertheless a relatively large amount per laborer. Usually, however, plants that have a heavy investment per laborer are large-scale.

There are a number of questions that require answers: (1) What determines the size of the fixed equipment? (2) What determines the amount of labor and materials that will be utilized with this fixed equip-

ment? (3) What determines the number of plants that will be brought under the management of a single firm? We shall consider each of these in the order named.

The size of the fixed equipment depends upon: (a) the extent of the market; (b) the advance of the arts in that particular industry, that is, inventions and technical processes; and (c) the supply of available capital. The modern giant flour mill would be impossible in a society in which the market was limited to (say) a hundred thousand people. Even though the inventions had been made and large-scale processes were well known, it would not be economical to build such a huge plant unless the market were large enough to utilize fully the fixed equipment. It was improvements in transportation, particularly the railroad, that widened the market and made possible large-scale production. Granted the availability of an extensive market, the size of the fixed plant will depend in part upon the technical conditions prevailing in the industry. The market for farm products is world-wide, and yet farms remain relatively small because the technical conditions in the industry are such that small-scale farming is more economical. The same is true of cigar-manufacturing, at least of the better-grade cigars. In some industries the technical conditions are such that only a large-scale plant can succeed at all. That is true, for example, in the steel industry. On the other hand, in the packing industry we have the huge packing plants at one extreme and the small country abattoir at the other. The large packing plant is based upon intensive division of labor and not upon technical processes requiring elaborate machinery. Finally, the large-scale industry is impossible in a country which is poor, and which has not a sufficient supply of capital to finance elaborate and expensive productive equipment.

The census of manufactures indicates that in many industries there is wide variation in the size of the establishments. This is partly due to the broad classifications used by the census. Thus all kinds of shoe factories are included under the heading "boots and shoes," those producing high-grade custom-made shoes as well as those producing cheap standardized shoes. The former are made in the smaller plants, the latter in large factories. But even where there are no such quality grades, there is sometimes wide variation in the size of the establishment. In some industries, however, small plants are nonexistent, being completely unable to compete. Table 12 gives the distribution of a number of industries as to size of establishments measured by the number of wage-earners employed. Only a few selected industries are given.

Group I includes industries in which large-scale establishments prevail. In these industries small establishments cannot survive. They are all industries that require much capital and a large investment in fixed plant. Group II includes industries that range all the way from small plants to huge establishments, but the large-scale establishments predominate. Small establishments survive, but they cut little figure in the total.

Group III includes industries that have all sorts of establishments, but the medium-sized establishment prevails. Group IV also includes industries that have establishments of all sizes, but the smaller establishments predominate. They are industries in which special machine processes of a large-scale sort are not greatly superior. Labor and management are more important than any special kind of capital equipment.

Group V includes those industries which have no establishments of huge size. Frequently they are industries which are tied down to local and restricted markets, or they are industries in which the technical processes are relatively simple, requiring little machinery, and in which labor or management or both are of much greater importance than capital.

We have discussed the conditions that determine the size of the fixed equipment. We turn now to our second problem. The amount of labor and materials that can profitably be employed in any marginal establishment will be that amount of labor and materials which will yield the lowest total cost per unit of output. This point depends fundamentally upon the law of diminishing returns, as explained in the first part of this chapter. Were it not for the fact that as more and more labor¹ is added to the fixed plant the point of diminishing product per unit of labor is reached, there would be no limit to the number of laborers that might be employed in any given plant. Because of the law of diminishing returns there is in every establishment a point of lowest total cost per unit of output, or it may be that there is a considerable range within which the total unit cost is approximately constant. This area or point of lowest total unit cost will generally be beyond the point of lowest unit cost in terms of labor alone. So long as the labor cost per unit of output rises less rapidly than the fixed capital costs per unit of output decline—and they will decline of course with every increase in out-

¹We are centering our attention here upon labor as the primary variable agent. There will of course be other variable agents also.

put—the total unit cost will continue to fall. But the point of lowest total unit cost will vary with the price paid for labor, as compared with the capital costs. If wages are relatively low, more laborers can be added before the point of lowest total unit cost is reached. Let us revert for the moment to Table 10. Suppose the cost of the variable agents is \$25 per unit instead of \$100; then the cost per unit of output in terms of the variable agents will be just one fourth as much as is given in the table. The total unit cost will then be lowest when 60 units of variable agents are employed instead of 50, as given in the table. If, however, wages become relatively high compared with capital and other fixed or overhead costs, the point of lowest total unit cost will be reached sooner and less labor can profitably be employed in any given establishment.

If, for any reason, the selling price of the commodity rises, more labor can profitably be employed. Thus during the war period, when prices were rising, establishments sought to increase output by hiring more labor. This was done even though there was no addition to the capital equipment. More labor could profitably be employed so long as the additional wage bill did not exceed the value of the product added by the more intensive utilization of the fixed equipment.

But the firm may continue to expand even though the plant has reached the most efficient size. This brings us to our third problem. The firm may build additional plants of equal size and efficiency. Certain advantages may follow from this procedure: (1) Freight charges may be reduced considerably by having plants located in different parts of the country. (2) Selling expenses may be reduced, for the greater the volume of sale the lower is the per unit cost of national advertising and of salesmanship. (3) A large firm usually can secure capital and credit more easily and on more favorable terms than a small one. (4) A large firm can frequently buy its materials on more advantageous terms. (5) A large firm can institute a single research establishment which will function for the several scattered plants. A striking example of this may be found in the General Motors Company, whose research and experimental department is utilized for all the General Motors cars: the Chevrolet, Pontiac, Oldsmobile, Buick, La Salle, and Cadillac.

There are limits, however, to the economical expansion of a firm in any line of business. Management in a large firm is necessarily distributed through a hierarchy of officials. Orders are passed down through various grades of subordinates. The organization becomes unwieldy and top-heavy. Red tape supersedes personal supervision and

attention. The large business unit faces at last the obstacle of the law of diminishing returns on entrepreneurship, and at last the point is reached at which it will be uneconomical to add more plants. Thus in most industries it is impossible for one business to expand until it occupies the whole field.

TABLE 12. Number of Wage-Earners, in Thousands

	SMALL ESTABLISHMENTS (1-20 WAGE-EARNERS)	MEDIUM ESTABLISHMENTS (21-100 WAGE-EARNERS)	LARGE ESTABLISHMENTS (101-500 WAGE-EARNERS)	VERY LARGE ESTABLISHMENTS (OVER 500 WAGE-EARNERS)
Group I				
Automobiles	0.6	5.0	23.4	181.6
Cotton goods	1.1	17.9	146.0	265.9
Glass	0.3	7.8	42.0	27.4
Steel work and rolling mills	0.2	4.8	55.0	315.1
Watches	0.0	0.2	1.3	14.4
Group II				
Agricultural implements .	1.6	5.2	14.7	32.9
Automobile bodies and parts	9.4	18.1	29.5	75.5
Boots and shoes	3.3	25.9	99.9	82.5
Railroad cars and repairs .	3.0	27.4	150.7	303.3
Slaughtering and meat-packing	4.7	11.1	13.3	121.8
Soap	1.2	1.9	4.8	12.5
Tobacco	0.7	2.3	5.2	10.1
Group III				
Chewing gum	0.3	0.8	1.3	0.8
Men's clothing	24.7	61.1	51.0	36.4
Confectionery and ice cream	19.0	27.0	34.3	15.5
Foundry and machine-shop products	42.0	116.7	191.8	132.3
Furniture	10.2	50.8	57.8	19.5
Knit goods	6.8	35.3	80.4	50.1
Cigars and cigarettes . .	19.6	28.3	52.6	38.2
Group IV				
Bread and bakery products	66.7	26.8	31.6	16.4
Carriages and wagons . .	5.9	5.7	4.8	1.8
Flour-mill and grist-mill products	20.8	13.4	8.1	3.2
Women's clothing	45.8	86.0	31.4	2.4
Jewelry	7.8	13.2	8.8	1.1
Group V				
Automobile-repairing . .	43.8	8.3	3.0	0.0
Butter	10.0	5.6	2.0	0.0

It must not be supposed that business men are always clearly aware of what constitutes the most efficient size of the business as a whole. That there is a most efficient size of business unit cannot be doubted, though it varies greatly in different industries and from time to time. It is a case of the survival of the fit. Many a business has been forced into bankruptcy by expanding beyond the limits of economy and efficiency. No doubt a business unit frequently continues to increase the volume of its output as long as the market is expanding. Yet such ill-considered growth will lead to trouble sooner or later. Such a firm is likely to find that smaller competitors with less top-heavy and unwieldy organizations are producing at lower cost. But it is not easy for businesses to stop growing when the most efficient size has been reached. The efficient concern is likely to grow; thus it may be big because it is efficient, and not vice versa. There is, however, danger of becoming too big. Hence businesses tend to pass from youth to maturity and thence to old age and final dissolution. Younger and more efficiently managed businesses take their places.

CONCLUSIONS

1. The law of diminishing returns affects (1) the entrepreneur's combination of employed agents, and (2) the size of the business.
2. The most economical combination is beyond the point of diminishing average returns.
3. The marginal entrepreneur will add variable agents until the lowest total cost per unit of output is reached.
4. The superior entrepreneur will add variable agents beyond the point of lowest total cost per unit of output.
5. All entrepreneurs will continue to add variable agents as long as these added agents more than pay for themselves.
6. The size of a single plant depends upon (1) the market, (2) current technical processes, (3) the supply of capital.
7. The size of the firm depends upon the number of plants that can be operated economically under the directive control of a single management. Here also the law of diminishing returns operates.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. In a period of intense business prosperity all firms tend to expand beyond the point of lowest total unit cost, but not beyond the point at which the marginal cost is equal to the selling price. Explain fully.

2. New businesses tend to be in a stage of decreasing cost; that is, an increase in output would result in lower total unit cost. Show why this is the case.

3. The railroad business as a rule tends to be in the stage of decreasing cost. What is the reason for this?

4. "We ought to force marginal entrepreneurs out of business in order that all our factors of production may be employed by superior entrepreneurs. Thus we should greatly increase the total national production." Criticize.

5. "The Ford Motor Company illustrates the point that there is no limit to the advantages to be gained from large-scale production." Criticize.

6. "The larger plants are bound to drive the smaller ones out of business sooner or later, and so in the end the tendency will be for one huge company to monopolize the field in each industry." Criticize.

7. "When the market is very small, no person can have any encouragement to dedicate himself entirely to one employment, for want of the power to exchange all that surplus part of the produce of his own labour, which is over and above his own consumption, for such part of the produce of other men's labour as he has occasion for."—ADAM SMITH. What bearing does this have upon large-scale production? What is the relation between division of labor and large-scale production?

BOOK TWO

=

VALUE

CHAPTER VI · Markets and the Marketing Function

DEFINITION AND CLASSIFICATION OF MARKETS

The term "market" is a familiar one the general meaning of which requires no explanation. If, however, we examine closely the usage of business and everyday conversation, we discover that the term includes three fairly distinct concepts: (1) A market may be thought of as a place where goods are bought and sold. (2) It is thought of also as the aggregation of interrelated organizations whose business it is to buy, sell, transport, and store a given sort of good. (3) Finally, a market is often spoken of as the sum of conditions or forces that play upon and determine prices. The first may be called the geographical, the second the structural, and the third the price-making (or "market-forces") concept of a market.

The geographical concept usually includes, in economic discussion, something more than the layman supposes. Wheat has a market in every hamlet where a wheat-buyer collects grain from farmers. But the geographical concept of the wheat market may include the entire world. Hence we frequently hear it said that Liverpool, which is the most important European port for buying and selling wheat, is a world market for wheat. This statement means that Liverpool, because of its importance, is the place on which the great dealers in wheat center their attention. Strictly speaking, however, there is no world market for wheat, but only a large number of related or linked markets which center in Liverpool. To prove this we have only to consider that wheat might sell at Liverpool on a certain day for the equivalent in American money of two dollars a bushel, while in other parts of the world a variety of prices were paid, because of interference with freedom of trade by tariffs and because of costs of transportation, storage, and cleaning, which must necessarily be incurred before the grain could be delivered in England.

The structural concept emphasizes the business organization of the market. A simple type of organization is exemplified by the markets for farm produce still found in many cities, in which the growers of veg-

etables and small fruits daily offer for sale the products of their gardens. Usually very little physical equipment is required for this market, merely a shed or rough building in which the sellers may display their products. To this place the local grocers, and frequently even consumers resort to lay in their daily supplies. There are no specialists, no accurate grading of products, and no service for keeping buyers and sellers informed as to the prices that are being paid for the same commodity in different parts of the building. Neither is there any service for reporting the amounts purchased at different prices each day and the amounts likely to be available on any succeeding day.

At the opposite extreme from the local farmers' market are the great produce exchanges for grain, cotton, and wool. In the United States wheat is purchased from the growers by a multitude of local buyers who clean, grade, store, and ship to larger buyers or warehousemen in the central markets, such as Kansas City, Chicago, and Minneapolis. The railroads carry the product and deliver it to the central markets, where it is purchased by millers, by speculators, or by large elevator companies. In these markets the sales are made in the grain exchanges between buyers and sellers who are specialists in dealing in wheat. Elaborate services are maintained for keeping both buyers and sellers informed as to the size of crops in all parts of the world, the amount of the visible—that is, readily available—supply, world demand, and current sales. These informational services are kept up by private agencies and by governments. The trading in any given exchange is so centralized that every trader can ascertain almost immediately how much of each grade of wheat is being bought as well as the price on each transaction.

The third concept of a market is much broader than either of the other two. It is also more abstract. The geographical area over which buyers and sellers are in close relationship is easily visualized. So, too, is the concept of the market as a group of interrelated business organizations dealing in the same commodity at the different stages of its progress from the primary producer to the consumer. But market forces are not always concrete things. They are conditions, facts, and relationships. Yet for the central problem of economics, price-making, this concept is more important than either of the other two. For example, when wheat sells for \$1 a bushel in Chicago on a certain day, and six months later for \$1.20, it would be futile to look to either the geography of the market or its structure for the explanation of the change in price. In practically every instance a variation of price is found to result from a variation of

one or more of the forces that control price-making. Perhaps the rise will have been caused by a report of unfavorable weather in Argentina, by a threat of war in Europe, or by better business conditions in the United States.

THE VARIABILITY OF PRICE-MAKING CONDITIONS

The conditions that affect the making of price are not the same for all commodities. On the demand side we find that some commodities are wanted for human consumption in their existing forms; others are wanted as aids in production. For some the consumer has a relatively stable demand, and for others his demand is fickle. On the supply side likewise there are a large number of different forces. The supplies of agricultural commodities are affected by weather, by the prevalence or absence of pests and plant diseases, by the money return that the farmers can get from other crops, by the supply of farm labor, and so on. The supply of steel, on the other hand, is not affected by conditions governing biologic growth, but it is affected by the cost of coal, the cost of transport, and other conditions that either do not affect the supply of wheat or affect it to a smaller extent. Both wheat and steel differ from certain consumable commodities in that the demand for them is not affected very directly by fad, whim, and the vagaries of demand that appear in the market for consumers' goods.

As a rule, commodities that can be easily graded and therefore easily described, such as wheat, can be sold over wide market areas, whereas those that cannot be so graded cannot be sold over so wide an area. The commodity sold over a wide market must be easily transportable, and it must be something that enters into trade in a multitude of ways, or that is widely consumed. In a wide market it pays those who deal in the commodity to specialize; and it is profitable likewise to build up systems of communication that put dealers over larger areas in touch with each other. Such markets are said to be well organized, and within them the goods dealt in tend to uniformity of price. For some goods, such as old dwellings, the market, outside large cities, is very poorly organized. Houses cannot be classified, and each buyer has a somewhat different notion of what constitutes desirable qualities.

Finally, we must distinguish markets in which both buyers and sellers are producers in the ordinary meaning of the word, or at least dealers, from markets in which only the sellers are producers or dealers and the buyers are consumers. The former are called *producers' markets* and

the latter *consumers' markets*. In the latter only consumable commodities are sold, whereas in the former the consumable commodities as well as raw materials and technical products are dealt in.

The contrast between producers' and consumers' markets is great even when we limit our examination to consumable goods. Producers' markets are usually better organized, the buyers have more competence to judge the quality of the things they buy, the force of competition is stronger, and the force of custom and habit is weaker than where the consumer is the buyer. The consumer buys for his personal satisfaction and for the satisfaction of his immediate family; his only criterion for judging whether a good is worth the price asked for it, when this price is common throughout the market, is his subjective estimate of the satisfaction he and his family are likely to derive from it. Although consumption, and not production, is rightly said to be the ultimate object of all economic activity, the art of choosing consumption goods intelligently and of utilizing them efficiently is, by contrast with business activities, extremely undeveloped. The buyer in the consumers' market is not an expert, and he cannot have the same knowledge either of the quality and serviceability of goods or of competitive prices as the dealer from whom he buys. In the producers' market, however, the buyer and the seller stand on a nearly equal footing in bargaining power and in knowledge of facts relevant to the transactions into which they enter.

Hence it is not surprising to find that competition works more efficiently in the producers' markets and that uniformity of price for goods of equal serviceability to the buyer is much more common than is uniformity of price for equal goods in the consumers' market. If one goes from shop to shop in the retail markets in a large city, he will usually find some variation in the prices at which even standardized articles are sold. And if he turns to articles whose demand is greatly affected by style, the variations become greater. To a certain extent each store has its own customers, its own special demand. Each is a somewhat separate market for goods plus the services that ordinarily go with them.

Yet the lack of price uniformity in the consumers' market must not be overstated. Each of these partially distinct markets is related to other markets. One store cannot go too far above the prices of its competitors, or it will lose its patronage. A "most common," or ruling, price for standardized goods prevails, and even for style goods the variations are not extremely large during the season when they are in

style. There are always some buyers who have the knowledge of quality, the time and inclination to shop, and the desire to find bargains, which enable them to ferret out the "best buys." They shift from store to store and spread the news of their bargains among their friends. Thus the seller is compelled to keep his prices nearly in line with those of his competitors. This fact is so well recognized that many merchants systematically "shop" competing stores.

The phenomenon of related markets is not confined to competition among the merchants in the same city. Investigations by Professor Secrist¹ and others have shown, for example, that the prices of clothing in different cities over a large territory are related. The consumer is not restricted to the stores in the city where he lives. He knows about prices and qualities in other places through personal observation, through reports of associates, and through advertising. If he believes that prices are too high or goods not in style in the local shops, he can and does go to other towns or cities to buy. And since any group of buyers may have the opportunity to buy in several cities, stores in different places tend to be connected or "linked" markets, in somewhat the same way that the wheat markets of the world are linked markets. But the connection is far looser in the retail markets, and in consequence the uniformity of price is far less.

It would be an error to suppose that the tendency of prices in different, but connected, consumers' markets to conform to a most common price is solely the result of the activities of the relatively small body of consumers who shop widely. Much of the uniformity that exists comes from the pressure of competition among the dealers themselves. A storekeeper is not compelled to remain always in one location or to keep to one line of merchandise. If prices rise in one place or in one line, and if profits rise with them, competition through the establishment of rivals is likely to put in an appearance. New shops will be set up by men just starting as independent entrepreneurs or by those who are leaving one line in the hope of bettering themselves in another. Many grocers have found in recent years that the maintenance of high prices or poor service in a given neighborhood creates a favorable situation for the establishment of a branch of a chain-store organization.

Yet in spite of the forces that tend to equalize the position of buyer and seller and to bring about uniformity of price, the operation of the price-making forces is different in the consumers' market from what

¹Horace Secrist, *The Widening Retail Market*. A. W. Shaw Company, Chicago.

it is in the producers' market. And to some extent the forces themselves are different. Convenience, pleasant surroundings and pleasant personal relations, custom, habit, and indifference to standards of efficiency are much more potent in consumers' than in producers' markets. Moreover, the consumer must buy for personal use and rely upon his own subjective estimates for testing the satisfactions he will derive from goods sold at different prices. The buyer in the producers' market is guided by more objective criteria.

MARKETING FUNCTIONS

In economic writings of a generation ago very little attention was paid to the subject of marketing. The reader was usually told that exchange was an important division of economics, that the objective value of goods was determined by exchange, that the people who engaged in bringing goods from farms and factories to consumers were productive laborers, and that those who took part in marketing were called by various names, such as "wholesaler," "broker," "retailer," and the like, to indicate that they performed different functions. It is only within recent years that the work done by marketing agencies has received the attention its importance warrants.

The present economic order is an exchange economy. No individual supplies himself with all the goods and services he consumes. Even the farmer, whose land often produces much of the food and fuel his family consumes, is compelled to sell to others a major portion of his products and to buy from others most of the consumable goods his standard of living requires. Among city dwellers self-sufficiency is still more impossible than it is on the farm. The manufacturer and his employees can use very little of their products, and professional workers require very little of their own services. All producers—farmers, manufacturers, factory hands, and professional workers—must rely on the aid of the people engaged in marketing to assist them in disposing of their products and to enable them to obtain the goods they consume. Accordingly, before we proceed further with the study of economic principles, it is very important to understand the functions performed by those who engage in marketing. Transport workers, merchants, and brokers play as important a role in production and price-making as do farmers and manufacturers. The following classification, given by Professor F. E. Clark, presents in compact form a general view of the functions performed by those who engage in marketing.

The Marketing Functions¹

- A. Functions of Exchange
 - 1. Demand creation (selling)
 - 2. Assembly (buying)
- B. Functions of Physical Supply
 - 3. Transportation
 - 4. Storage
- C. Auxiliary or Facilitating Functions
 - 5. Financing
 - 6. Risk-taking
 - 7. Standardization

The function of demand creation involves the work of finding buyers for the products of the farmer and the manufacturer. Goods do not sell themselves, any more than plants grow without cultivation or factories run without human direction and control. Even such staples as elementary raw materials and foods must be brought to the attention of the buyers, who must be informed where the articles are to be had, what their qualities are, and at what prices they are offered. Luxuries and semiluxuries, the vast range of things used nowadays without any reference to physical necessity, require more effort and expense for selling. Not only must they be called to the attention of the buyers, but their uses must be explained, and either by advertising or by more direct methods the buyer must be convinced that they will satisfy a particular want. Sometimes these sales methods work to the detriment of the consumer, who may be persuaded to buy something which he finds later does not satisfy his needs or which is inferior to some other good he might have purchased at the same price. On the other hand, the selling function unquestionably extends the range of consumption and increases the satisfactions of the consuming public.

The function of assembly does not mean the work of gathering together the physical items in a stock of goods, for that is transportation. "Assembly" means that the middleman acquires the legal title to the goods, or at least the right to buy or sell them. He may have at his disposal, for example, a million bushels of wheat. A part he may actually own and have stored in various places. Another part he may not own at all; he may merely have agreements with other dealers to deliver it to him at a future date. But in either case he can control that amount, can decide to sell it or withhold it from the market. He puts himself in a position to sell large or small amounts as the buyer may

¹ F. E. Clark, *Principles of Marketing*, p. 11. Copyright, 1922, by The Macmillan Company. Reprinted by permission.

wish, and he prepares to sell different grades to different buyers. In the marketing of manufactured goods the function of assembly includes a slightly different sort of work. The department store must stand ready to sell to its patrons literally hundreds of different wares, very few of which can be obtained from a single manufacturer. Someone must find out where all these wares can be obtained, the quantities in which they can be obtained, and their prices. Then he must offer them for sale in definite amounts. All this the department store does for its patrons. The smaller store, which cannot afford to perform this function, usually relies on a wholesaler or some other intermediary.

The function of physical supply, which includes storage and transportation, is almost self-explanatory. Everyone knows that most of the articles he consumes are not produced at his door and that they must be brought from more or less distant places before he can consume them. He knows also that many products, such as agricultural crops, are harvested only at certain seasons and that they must be held in storage so that consumption may be continuous. Manufacturing is not ordinarily so much controlled by the seasons as is agriculture, but the demand for manufactured products may be seasonal. Hence storage must be provided if production is to be continuous. Furthermore, both production and consumption may be continuous and storage still be necessary because a stock must be maintained to give the consumer a choice among the goods he buys. The merchant is compelled to keep a stock on hand also because he cannot foresee exactly what the consumer desires. Often the latter makes his decision by looking over the different styles or varieties of goods displayed. Finally, local storage is sometimes resorted to when it is cheaper to ship goods to a particular city in large amounts and then store them until gradually they are consumed than it is to ship in relatively smaller amounts and avoid the major part of the expense of storage.

The auxiliary functions are so named because they do not assist directly either in the physical supply of goods or in the finding of buyers for them. Financing is a necessary function because time must elapse between the completion of the physical process of production and the sale of the goods to the consumers. The manufacturer and the farmer or some middleman must have money tied up in the stocks waiting to be consumed. The interest on these funds constitutes the financing service for which someone must be paid. Risk in marketing arises from (1) the possibility of an adverse change in price while the goods are in the hands of the dealer, owing to a falling off of demand or to over-

supply, (2) the possibility of physical deterioration, and (3) the risks arising from a failure of buyers to live up to their agreements, either their failure to take goods ordered or their failure to pay for them. For many sorts of goods, particularly things for personal use, the risk of an adverse change in price is great. The style may change before the merchant has sold his stock, or the fad for a particular kind of good may not appear in one city even when it has appeared in other places. Or the season may be unusual, and the seasonal goods that the producer and the marketing agencies had reasonably anticipated would be in demand fail to attract buyers, as is true of men's straw hats when the summer is cool and rainy. All these risks must be borne by the marketing agencies; and because the existence of a risk implies that there will be losses, the goods actually sold must bear the cost of those remaining unsold or sold at very low prices.

Standardization of products, which means classification into established groups or grades, is an important marketing function for some kinds of products. When the buyers and sellers cannot come together with the goods before them, grading becomes extremely important. At present wheat is sold by description, that is, by grade, all over the world. A miller in England can buy "No. 1 dark northern spring wheat" and know approximately what he will receive, because wheat has been graded and standardized. Many commodities cannot be standardized, because technical tests for certain qualities have not been devised or because the make-up of the good is constantly changing. For both reasons it appears impossible to standardize style goods—clothing, millinery, shoes, hats, and so on.

MARKETING AND SPECIALIZATION

In the stages of economic development that preceded the present highly specialized economy each producer supplied a relatively limited area and relatively few consumers. Lack of cheap transportation, the slowness of communication, the simplicity of production, the absence of large plants, and political conditions all contributed to keeping the products of farmers and craftsmen from moving very far from the places where they originated. The producer could readily ascertain the kind and quantity of product his patrons required, and the consumer could more easily than now hold the producer responsible for poor quality and misrepresentation. By reason of the influence of custom and the close adjustment of production to needs, prices were much

more stable. There were, of course, exceptions to the conditions just sketched. A few commodities were imported into every country at extremely high costs for transportation, financing, and risk-bearing; but local production for local consumption was the rule.

Today conditions are very different. The average laborer's family consumes in the course of a year things that have come from literally every part of the globe: coffee, sugar, rice, rubber, and so forth. Even the products grown and manufactured within his own country are derived from the factories and farms of people he has never seen and knows nothing about, and most of these products have passed through the hands of many dealers, after leaving the factory or the farm, before they are delivered to him. Marketing has come to play a vastly more important role. And with this development some of the advantages of the old order have been lost. The relative cost of marketing has doubtless become greater in comparison with the cost of physical production, that is, the growing, mining, and processing of goods. In Table 13 are given estimates of the cost of selling at retail five important types of consumers' goods.

TABLE 13. Cost of Operating Retail Stores, by Percentages of Each Dollar of Sales¹

	1913	1916	1918	1920	1921
Clothing	25.90	25.20	26.50	25.70	24.03
Hardware	20.13	18.60	18.52	19.50	24.90
Shoes	24.14	24.45	24.07	25.25	26.88
Dry goods	26.90	26.30	27.30	26.00	no data
Groceries	14.70	15.20	14.30	14.60	16.80

The percentages given in the table above show only the *retail* cost of marketing, and not the proportion of the consumer's dollar that goes to pay for all the marketing services performed in respect to the five types of commodity. A quantity of groceries, for example, that cost the retailer a dollar was bought by the wholesaler for about ninety cents during the years for which the table shows estimates.² The manufacturer of these foods also had incurred some selling expense, for which he was paid, and the dealers from whom he bought his raw materials were also compensated for their services. The con-

¹ Report of the Joint Commission of Agricultural Inquiry, Part IV (Marketing and Distribution), p. 176.

² *Ibid.* p. 158.

gressional commission from whose report Table 13 was taken is authority for the statement that "it costs more to sell and deliver the marketable products into which agricultural commodities are converted than to produce and manufacture them."¹ This estimate seems, on general observation, somewhat higher than would be true for all goods, but we must remember that many of the goods and services for which we spend our money incomes are not subject to such heavy marketing expense. Services rendered by the government, such as sanitation, education, and recreation, have few or no marketing costs. Personal services furnished by physicians, household servants, caterers, barbers, and the like are not marketed in the proper meaning of the term, and therefore their prices include little or no expense for that service.

In the case of most commodities the function of marketing is more or less combined with that of manufacture. It is therefore difficult to estimate what the cost of marketing really is. All that can be accurately stated is the relative amounts of the consumer's dollar that go to the producer of the materials and to other groups. The data in Table 14 show the proportions for certain important agricultural commodities.

Another way of showing the importance of the cost of marketing is to compare the number of persons engaged in trade and transportation with the total number of gainfully employed in the United States. In 1930, 48,830,000 persons were gainfully employed. Of these, 6,081,000 were employed in trade and 3,843,000 in transportation. If we exclude from the classification of trade such doubtful groups as bankers and bank employees, real-estate agents, undertakers, and insurance agents and officials, there still remain 5,300,000, or 10.9 per cent of the total of gainfully employed, who were chiefly engaged in marketing commodities.²

Still another method of showing the cost of the present marketing system in general terms is to estimate the proportion of the national income received by those who derive their incomes from marketing. According to W. I. King, the total realized income of the United States in 1928 was \$89,419,000,000. Of this total, \$13,137,000,000, or 14.7 per cent, was received by the people in mercantile occupations. In 1925, \$6,736,000,000, or 7.6 per cent of the national income, was received by those engaged in transportation.³ Since a part of the business of trans-

¹ *Ibid.* p. 115.

² Statistical Abstract of the United States, 1932.

³ W. I. King, *The National Income and Its Purchasing Power*, p. 94. National Bureau of Economic Research, 1930.

TABLE 14. Estimated Retail Value and Equivalent Farm Value of Quantities of Food Purchased by a Typical American Working-man's Family¹

YEAR	FARM VALUE (IN DOLLARS)	RETAIL VALUE (IN DOLLARS)	MARGIN (IN DOLLARS)	FARM VALUE AS PERCENTAGE OF RETAIL VALUE
1913	134	252	118	53
1914	137	258	121	53
1915	134	258	124	52
1916	155	285	130	54
1917	223	370	147	60
1918	245	424	179	58
1919	267	470	203	57
1920	272	514	242	53
1921	179	404	225	44
1922	170	374	204	45
1923	173	384	211	45
1924	170	381	211	45
1925	198	410	212	48
1926	202	418	216	48
1927	190	406	216	47
1928	194	407	213	48
1929	195	415	220	47
1930	171	391	220	44
1931	121	322	201	38
1932	88	270	182	33
1933	92	264	172	35
1934	108	295	187	37
1935	138	331	193	42

portation consists in carrying people and not goods, by no means all the income received by this occupational group can be counted as a cost of marketing. On the other hand, not all the persons concerned in marketing are included within the two groups for which the statistics have been given. All farmers and some of the workers connected with every manufacturing establishment assist to some extent in marketing, and the incomes which these groups receive on account of that work is a part of the cost of the nation's marketing service.

Against these costs that attach to the present system of marketing and have probably become proportionately larger as markets have widened and grown more complex must be set very important gains.

¹*Price Spreads between the Farmer and the Consumer.* United States Department of Agriculture, 1936. Mimeographed. No allowance made for the processing taxes, amounting to about \$2 in 1933, \$10 in 1934, and \$11 in 1935.

Although no statistical data are available to prove the point, there can be no doubt that the factory system, specialization, and large-scale production, all of which depend on wide markets for their existence, have increased tremendously the average productivity and income of the population of the United States and western Europe during the last century and a half. And if we should discard the present arrangements, which are admittedly expensive, and revert to the simpler market organization prevailing in the early eighteenth century, we should lose a large portion of our present income. Without the present system many kinds of goods could not be produced at all, and others only at higher costs.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "If the chain of middlemen that connects the producer with the consumer were unnecessary, it would never have come into existence." Is this statement correct? Does it contain complete proof that present arrangements are the best that can be devised?

2. It is sometimes said that advertising and other forms of demand creation are entirely wasteful. Like modern armaments, they result from destructive competition. If dealers did not attempt to sell goods, how would the consumer find new goods? Would he be able to discover as easily as now the cheapest place to buy? If it is assumed that advertising is purely competitive, is it socially wasteful?

3. In which kind of market should you expect prices to be more nearly uniform, producers' or consumers'? Why?

4. "Instead of having highly centralized livestock markets in Chicago, South St. Paul, Omaha, and other places, we ought to have a packing plant within a radius of fifty miles of every farmer so that each farmer could sell his livestock directly to the packers." Discuss.

5. "A farmers' and wage-earners' co-operative marketing association ought to be organized so that farmers could sell their products directly to the consumers in the cities. In this way farmers would get more for their products and wage-earners would pay lower prices for foodstuffs." Discuss.

6. "These Utopian reformers want to accomplish the impossible. They want everyone to sell dear, and at the same time they want everyone to buy cheap. They want a high price for beef on the hoof and a low price for beef on the table." Discuss.

CHAPTER VII · The Price-Making Process: Demand and Supply



PRICE AND VALUE

The value of any commodity is its power to command other goods in exchange. This value is measured by the quantity of the other goods for which the commodity in question exchanges. If shoes are bartered for cloth at the rate of one pair of shoes for six yards of cloth, then the value of cloth is measured in terms of shoes, and the value of shoes in terms of cloth. This is a correct representation of power in exchange; but it is not of much significance, because it only states the exchange value of each good in terms of one other good. If we wished to know the general exchange value of shoes, we should have to discover the rates at which they exchanged for all the other goods for which they might be bartered. To find these exchange relations would be practically impossible. And when exchanges were carried on by barter, it is hardly conceivable that anyone had very definite notions of the value of any good in terms of more than a few other commodities.

In present-day industrial societies barter has been superseded by the exchange of goods for money. The amount of money for which a unit of any given good exchanges is called its *price*. Since exchanges are now commonly made in money, it is possible to find the value of any commodity by comparing its price with the prices of a large number of other goods. If it is found that shoes sell for six dollars a pair and cloth for one dollar a yard, then it is easy to show that the exchange value of cloth and shoes at current prices is six yards of cloth for one pair of shoes. Comparisons of the price of cloth or shoes with the prices of many other goods would establish a still more complete measurement of the exchange value of both. And this comparison is constantly being made in many places and by many persons. Hence we infer that the price of a commodity is a more or less accurate measure of its exchange value.

Price is not, however, an entirely accurate measure of value, because, as will be shown later, the value of money, which serves as a basis for the computation of the values of other goods, itself may change. But,

for the purpose of simplifying the analysis of the forces that control value, we can assume that the value of money is constant and that if any good rises or falls in price this rise or fall reflects a change in its value.

THE ANALYSIS OF PRICE

If a person should attempt to find out why a specific commodity sold for a certain price at a particular time, he would be confronted at once by a bewildering mass of facts whose relations are extremely complex. Moreover, if he tried to measure the relation of each fact to every other, he would never arrive at any conclusion, because the human mind is not capable of dealing with a very large number of facts or relationships at one time. Since he could not deal with the large number of facts and all their interrelationships, in all probability he would try to classify many of the facts into a limited number of groups, and to each group he would then attach a name. Thereafter, when any fact relevant to his investigation was found, it would be assigned immediately to its proper group. This procedure, which is called classification, is common to all scientific investigation.

When different facts are lumped together, it is evident that some things will be put into the same group and treated as though they were exactly alike when, in reality, they are more or less different. This involves a certain amount of error, not only in the classification itself, but also in any statement concerning the relationships between groups of facts. Hence the common saying that any generalization, that is, any statement concerning the relation of one group of facts to another, is partly false. The greater the differences between the facts that are put into the same group the greater is the probability that any general statement about their causes and effects will be partly untrue. The danger of making erroneous statements because of the classifying of unlike things together is a very real one. It would be an error in reasoning, in many cases, to conclude that because the average wages of factory workers in some community had risen, the wages of a particular group of workers had also risen. Or again, when the general level of prices has risen 10 per cent, owing to a decline in the value of money, it does not follow that the price of each kind of good has risen by the same percentage.

Still another useful procedure in dealing with very complex relationships between facts is to rule out of consideration in the beginning all but the most important factors that appear to produce a given result.

This leads to piecemeal analysis. The student may say to himself: Let us see how any given price would be affected by a few of the conditions that appear to control it. This procedure is called abstraction. Like classification, the method of abstraction is replete with opportunities for making false steps in reasoning. It might be concluded, for example, that if a larger crop of cotton were harvested, the price would be lower. But this deduction assumes that "other things are equal"; that is, that the demand of consumers for cotton cloth has not changed, that there have been no changes in the competition among the mills that manufacture cotton, that the growers will not be able to hold back part of the crop for future sales, and that the amount of cotton and cotton cloth remaining from last year's output is normal.

In the analysis of price we shall proceed in the following manner: We shall begin with the examination of what is called in economics "market price" and attempt to understand the way in which economic forces control prices for short periods of time. Having explained how prices are determined for a short period of time, we shall then go on to examine the forces that determine prices over a considerable period of time. This leads to the analysis of normal price.

MARKET PRICE

It is commonly said that market price is the price for which a good actually sells in any given market at any particular time. Yet in actual practice economists or market analysts seldom attempt to explain *particular prices*. The forces which cause a good to sell now at a little higher and later at a little lower price, and which give rise to minor differences in the price of the same good in the same market, are numerous and intangible. Moreover, small mistakes in judgment by buyers and sellers are too frequent to permit the explanation of all the minor changes and discrepancies actually encountered in the market. The market price we shall consider is that which is most common, or typical. If the price at which the bulk of the transactions in wheat take place during a certain day or week is one dollar, then we shall assume that this is the market price for wheat during that period. Many sales will have been made at somewhat higher or lower prices, but with these deviations we shall not be concerned. The justification of this procedure is the fact that for the analysis of price-making forces certain facts are not significant, because they are unlikely either to appear very often or to have very great effect when they do appear.

Market price is commonly contrasted with normal price. The latter is defined as the price that would prevail if only the more permanent forces affecting price-making were operative and if sufficient time were allowed to permit these forces to bring about a completely stable equilibrium of demand and supply. At this point we need not attempt an explanation of normal price, but the contrast between market and normal prices must be grasped clearly if the analysis of market price is to be understood. The analysis of market price differs from the analysis of normal price by taking into account somewhat different forces and by considering the short-run effects rather than the long-run effects of all forces. Cost of production, for example, is not a considerable factor in the determination of market price because, as we shall see later, it operates slowly, and in the short-time fluctuations of market price it could be only a minor factor; but it is one of the most important in determining normal price. Fluctuations in the demands of buyers are exceedingly important in the analysis of market price but of less significance in the determination of normal price, because the passage of time allows producers to make adjustments in supply that largely, though not wholly, offset the changes in demand.

DEMAND AND DEMAND PRICE

The simplest statement concerning the determination of market price is that demand and supply control it. This is the explanation most frequently heard in the market place and on the street. And the statement is true. For as the terms "demand" and "supply" are commonly used, they are taken to include all the forces that influence buyers and sellers. But like most other general and all-inclusive explanations, this formula is of very little assistance in understanding the actual phenomena of price until the terms "demand" and "supply" have been exactly defined and until the complex forces for which they are but convenient names have been analyzed. Failure to observe these scientific procedures leads to many hasty and incorrect generalizations. For example, we are sometimes told that trade-unions cannot permanently better the economic conditions of the working classes because "wages are determined by the immutable law of supply and demand"; or we read that because of the appearance of a monopoly "the law of supply and demand has been repealed"; or we are informed that "price having risen, the demand for the good fell off." Now as a matter of fact, the principle of demand and supply does not preclude the successful establishment and

operation of collective bargaining among laborers or of effective monopoly among producers. Neither is it true that demand falls off as a result of the rise in price. These and many other erroneous applications of the principle of demand and supply only illustrate the necessity for proceeding slowly and carefully in the analysis of the price-making process.

The first step in the analysis is the definition of demand. *Individual demand is the amount of any given good an individual stands ready to buy at a particular price, during a given time, and at a given place.* Demand is the amount the buyer *stands ready to buy* or, more briefly, *will buy*, and not the amount he has taken in the past. His demand cannot be measured exactly by computing the amount of the good sold him in the past, for this depends, not on demand alone, but also on supply. This distinction is of great importance. In the first place, one may demand a good without making a purchase, for if the price he is willing to pay for the smallest obtainable quantity is below the figure at which the merchant is willing to sell, no sale can be made; yet his demand does unquestionably exist, and if the conditions of supply should change, the buyer who is now excluded might be able to buy. These excluded buyers are often of considerable importance; a producer who can secure the economies of large-scale production must frequently rely upon them to increase the volume of sales when prices are reduced, and partly because of them the output of plants can be sufficiently enlarged to permit these economies to take place. Second, the amount the buyer has purchased in the past is not an exact guide to the amount he will buy at some future time, even though the price the sellers ask remains unchanged; his needs or desires may have altered, or his income may have become larger or smaller. Third, it is logically incorrect to speak of demand as being the amount of a good purchased at a given price, since price and the volume of sales are the results of both demand and supply. The price at which the sellers hold a good influences the volume of sale.

If demand is the amount of a good which the buyer stands ready to take at a given price, then the *individual-demand price* is the quantity of money which the buyer will give for a definite amount of a good. Price and amount mutually condition each other. For this analysis the price must be stated before the amount the buyer stands ready to take has any significance, and the amount must be known before the price means anything. A consumer might be willing to pay a dollar a dozen for fresh eggs, but in all probability he would stand ready to take only a very limited quantity at that figure. Or he might be willing to buy two dozen a week, but not to pay a dollar a dozen for them.

The two constituent elements of demand are desire and ability to pay. Desire is fundamental, because where it does not exist people will not buy at any price. On the other hand, ability to pay is also essential, because without it demand cannot be made effective. When goods sell at low prices, the common conclusion is that these low prices are primarily due to the failure of the goods to awaken an intense desire on the part of the consumer. This conclusion is frequently correct, but it is not always so. Sometimes low prices result from the lack of purchasing power. Striking examples of the effect of a reduction of the purchasing power of consumers upon the prices of many sorts of goods are found in every severe business depression. In the period from 1930 to 1933 the retail prices of necessities of life fell precipitously. The explanation of these declines was the fact that many people had had their incomes greatly reduced and dealers in bread, milk, potatoes, meats, and many sorts of clothing could sell the amounts of their wares they wished to sell only by taking lower prices for them. Some luxuries are demanded in quantities and at prices that seem entirely out of proportion to the objective usefulness inherent in them. Usually the explanation is that they are bought by the rich and the well-to-do, who are able to pay handsomely even for things for which they have no very intense desire.

THE INDIVIDUAL-DEMAND SCHEDULE

It is commonly observed that an individual buys less freely of any commodity when the price of that commodity is high than when it is low. Speaking generally, we find that when the price of any good falls, the volume of purchases by the individual consumer increases and that it decreases when the price rises. This relationship between price and the volume, or amount, of purchases gives rise to the individual-demand schedule, which is a list of the quantities of any given commodity the individual stands ready to buy at varying prices.

Illustrative demand schedules of one buyer for two commodities are given in Table 15. The table shows that at 65 cents the buyer refuses to consider the purchase of either; that at 60 cents he stands ready to buy sparingly of A, but that not until a price of 40 cents is reached will he take any of B; and that at relatively low prices he is prepared to buy both commodities in large amounts. It must not be supposed that the average consumer is necessarily conscious of the different prices he would pay for different amounts of any commodity. If he has been accustomed to buy milk at 12 cents a quart, he is probably quite uncer-

TABLE 15. Individual-Demand Schedules

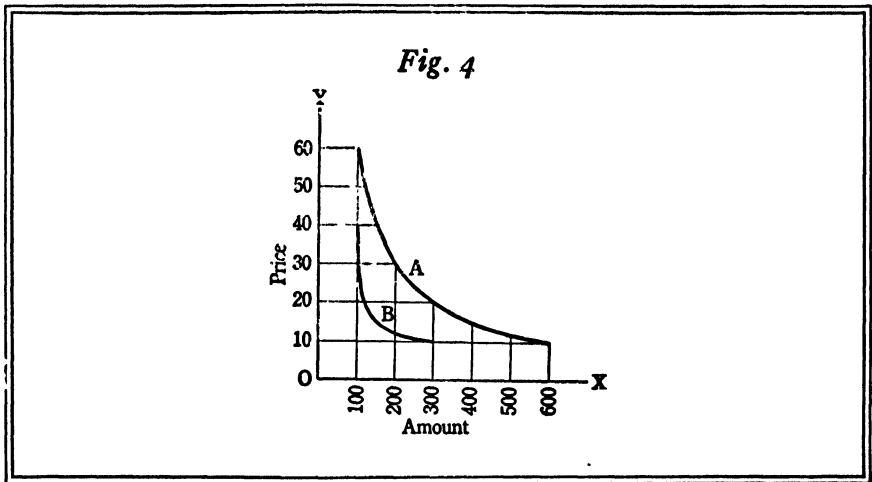
PRICE PER UNIT (IN CENTS)	AMOUNT BUYER WILL TAKE	
	Commodity A	Commodity B
65		
60	100	
55	109	
50	120	
45	133	
40	150	100
35	171	102
30	200	105
25	240	110
20	300	120
15	400	150
10	600	300

tain as to the amount of it he would be willing to buy at 18 or at 6 cents a quart. But the existence of a demand schedule does not require that he should be in possession of a plan of purchase for all probable prices. Studies of the marketing of goods show that even though consumers may not be conscious of the course of action they would take if the price of a commodity changed, purchases do decrease when the price of the commodity rises and increase when it falls, if incomes remain unchanged.

DEMAND CURVES

When two economic quantities vary in relation to each other, it is often convenient to show this relationship by means of curves. In Fig. 4 the relationship between the prices and amounts of commodities A and B (Table 15) are thus illustrated. Two lines, *OX* and *OY*, are drawn to intersect at right angles at *O*. Price is conventionally measured by intervals on *OY* and amounts by intervals on *OX*. In plotting the two demand schedules given above, the procedure is as follows: First, *OY* and *OX* are laid off at convenient intervals. (In this illustration price intervals of 10 cents and amount intervals of 100 units are convenient for the entire scale.) Beginning with commodity A, we find that 100 units is the smallest amount given in the schedule and that the corresponding price is 60 cents. So we insert a dot directly above the 100-mark on *OX* and directly opposite and to the right of the 60-mark on *OY*. Next, we place a dot directly above the 200-mark and opposite the 30-mark, and so on until the final entry in the demand schedule is

represented by a dot above the 600-mark on OX and opposite the 10-mark on OY . The next step is to draw a line connecting all the points, which is called the demand curve for commodity A. By similar procedure the points in the demand schedule for commodity B are determined and the curve representing that schedule drawn in. At 10 cents our hypothetical individual ceases to increase his purchases of A; and if the price is above 60 cents, he stops buying altogether. Hence we have terminated the curve A at these points.



The advantage of translating demand schedules into demand curves, as well as certain dangers in the procedure, must be observed at this point. From the schedule it is seen that when the price falls from 20 cents to 15 cents, the amount of A that will be taken increases at the same time from 300 units to 400 units. As far as the schedule shows, we know nothing about the amounts that will be taken at prices between 20 cents and 15 cents; but when we draw a curve, we do actually represent the amounts that will probably correspond to the smaller price intervals. We assume that if the amount increases at a certain rate when the price falls from 20 cents to 15 cents, its increase between these price points is regular. Since the amount interval in the schedule corresponding to the price interval of 5 cents is 100, we assume that for every decline of 1 cent in price the amount increases about 20 units. Where the entries in the demand schedule, either price or amount, are far apart, the curve enables us to gain some notion as to the amounts that will be taken at intervening price points.

Suppose that we wish to find the amount of A that the buyer will take

at $17\frac{1}{2}$ cents. The first step is to ascertain the point on the demand curve, *A*, that corresponds to the price of $17\frac{1}{2}$ cents on *OY*. Then the point on *OX* directly below this point is ascertained. In this case we find that a price of $17\frac{1}{2}$ cents corresponds approximately to an amount of 350 on the axis *OX*. We therefore infer that although the schedule does not show an entry of any amount opposite $17\frac{1}{2}$, the buyer will probably take 350 units at that price. In coming to this conclusion we have, however, assumed that if the price declines from 20 cents by 1-cent intervals, amounts will increase by intervals of 20 units. As far as the schedule of demand for *A* is concerned, this may or may not be true. For large numbers of consumers' goods, such as common kinds of food, for many minor items of clothing, and for a multitude of small articles our assumption of continuous and regular variations of amounts between the two price points is essentially correct. For these goods the curve represents the actual conditions with a considerable degree of accuracy. It is well established, however, that the curve does not represent the facts in respect to articles purchased only occasionally and in relatively large, costly units. A consumer may buy a pair of shoes each season of the year at \$10 a pair and decline to buy an additional pair at any price above \$5 or \$6. Or a family may install a telephone in their house at a rental of \$3 a month and refuse to have a second put in even though the company changes its rates and offers to install and operate at \$2 a month. Of course it might well happen that if the household were conducted on a more elaborate scale, a second connection would be taken at only a small reduction in price. To sum up: For some commodities a curve does not represent the demand schedule of the individual, because the variation in amounts he stands ready to take is not continuous but moves by large jumps instead of creeping by small intervals. The correct representation for such a demand schedule is a series of crosses or dots to show the variation. For other commodities there is continuity, and small price changes will be accompanied by small variations in amount.

COLLECTIVE DEMAND

Among the consumers who patronize the same market there are many who demand the same good. Taken together, they make up the collective demand for it, and the amounts they stand ready to buy at different prices make up the collective-demand schedule. There is, of course, much similarity between the demand schedules of a large number of individuals buying the same commodity in the same mar-

TABLE 16. Individual-Demand and Collective-Demand Schedules

PRICE PER UNIT (IN CENTS)	I INDIVIDUAL SCHEDULE (AMOUNT)	II INDIVIDUAL SCHEDULE (AMOUNT)	III INDIVIDUAL SCHEDULE (AMOUNT)	IV COLLECTIVE SCHEDULE (AMOUNT)
60	80	20		100
50	90	30		120
40	100	40		140
30	110	50	10	170
20	130	70	20	220
10	160	100	50	310

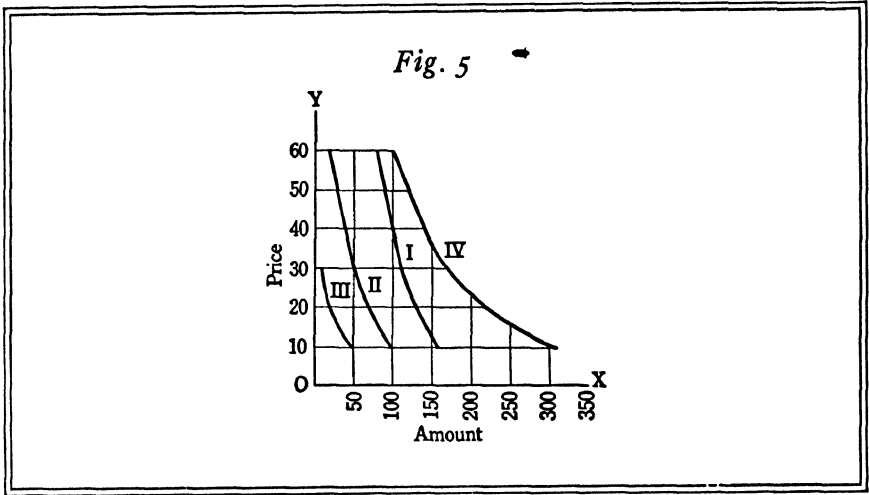
ket, and yet there is always some degree of dissimilarity. There are differences in tastes, family requirements (because of differences in size, age-composition, and standards of consumption of the different families), family incomes, and occupations. Hence variations in the collective-demand schedule are not identical with variations in the demand schedule of any individual except by chance. Thus, in the preceding example we found that the individual would increase his buying by one third if the price fell from 20 cents to 15 cents, but it is improbable that the same change in price would affect the buying of the entire consuming public in exactly the same proportions. This lack of similarity between a collective-demand schedule and any of its component individual schedules is illustrated by Table 16, in which three individual-demand schedules are combined to make a collective-demand schedule. These schedules are represented graphically in Fig. 5 on page 104. A comparison of the demand curves in that figure shows that the collective-demand curve (*IV*) is not the same in slope or elasticity as any of its component parts.

CHARACTERISTICS OF DEMAND SCHEDULES

If we investigate the amounts of different commodities sold within any market, we find that there is the greatest variation. Some articles are sold in very large amounts, and others are purchased only in very limited quantities. Some can be sold only if their prices are relatively low, and others, although they are taken in larger amounts when their prices fall, can be disposed of in large quantities even when their prices are relatively high. From these facts we conclude that the demand schedules of the individuals who make up the entire body of buyers are unlike in respect to different goods. Referring again to Fig. 4, page 101,

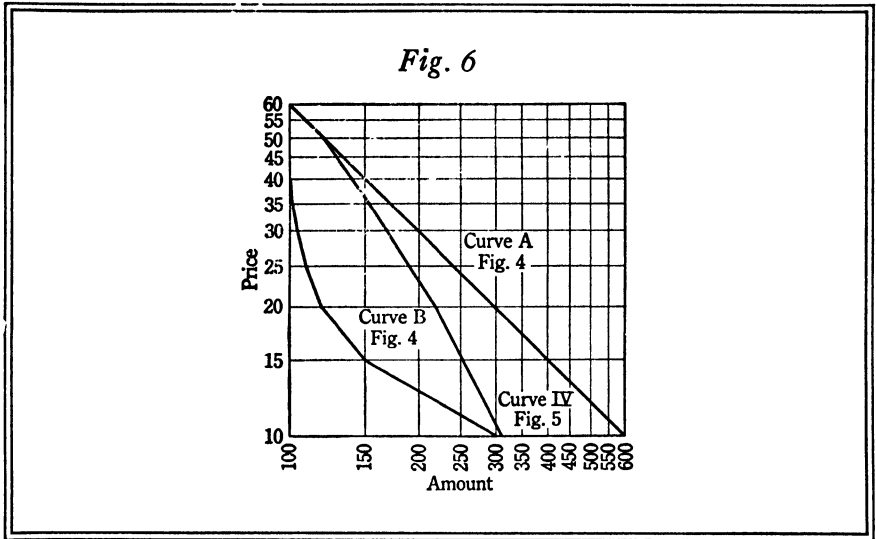
and the demand schedules in Table 15, we find that curve *A* differs from curve *B* in several important ways. It has greater extent; that is, it shows that the buyer will take the good through a greater range of prices. And it has greater amplitude than the other curve; that is, it shows that the buyer will take larger amounts at given prices.

Furthermore the demand schedule represented by curve *A* has greater elasticity at some prices and less at others than that represented by



curve *B*. The demand for a commodity is said to be *elastic* when a relatively small change in price is accompanied by a relatively large change in the amount the buyer (or the entire market) stands ready to take. If, on the other hand, this change in price is accompanied by a relatively small change in the amount the buyer stands ready to take, the demand is said to be *inelastic*. For example, if a drop of 1 per cent in price is followed by a much larger percentage increase in the volume of sales, the demand is said to be very elastic; but if a 1 per cent decrease is followed by an increase in volume of sales of much less than 1 per cent, the demand is very inelastic. If the rate of increase in sales exactly equals the rate of change in price, the demand is said to have an *elasticity of unity*. In terms of Fig. 4, the demand for commodity *A* has an elasticity of unity throughout its entire extent, and the demand for *B* has an elasticity of less than unity between 40 cents and 15 cents, and an elasticity of more than unity between 15 cents and 10 cents. If it became parallel to *OX*, it would have an elasticity of infinity, and if parallel to *OY*, of zero.

The elasticity of demand cannot be tested by determining whether the *absolute* increase in sales is great when price falls a certain amount or by determining whether the decrease is absolutely large when price rises. In Fig. 5 the demand represented by Schedule II is relatively elastic between the prices 60 cents and 40 cents, although the absolute increase in amount demanded is only 20 units. The demand represented by Schedule I is *relatively* inelastic between these prices in spite of the



fact that the *absolute* increase is exactly the same as in Schedule II. In Fig. 4 the elasticity of demand for commodity A is about the same between the prices 15 cents and 10 cents as it is between 60 cents and 55 cents, although the absolute increase in amount demanded is 200 units for the former price interval and 10 units for the latter. The change from 60 cents to 55 cents is *relatively* small, and it needs only to be accompanied by a *relatively* small change in amount to conform to the definition of elasticity of unity. On the other hand, the change in price from 15 cents to 10 cents is 50 per cent of the lower price and must be accompanied by a proportionately large increase in the amount to agree with the definition of elasticity of unity.

It is evident that one can judge the elasticity of demand from curves plotted on an absolute scale, as in Figs. 4 and 5, only with the greatest difficulty, because an ordinary arithmetic scale does not show *proportional change*. In Fig. 6 the two demand schedules, which are shown on absolute scales in Fig. 4, have been plotted on logarithmic scales, which

are arranged to show proportional variations of both price and amount; that is, the intervals on the amount scale (corresponding to OX in Fig. 4) are so spaced that the distance on the chart between 100 and 200 is exactly the same as between 200 and 400 (since 100 is to 200 as 200 is to 400). On the absolute scale the interval between 200 and 400 is twice as great as between 100 and 200. The intervals on the price scale in Fig. 6 (corresponding to OY in Fig. 4) are likewise so constructed that a given linear distance represents the same proportional change; for example, the distance between 10 cents and 20 cents is the same as between 20 cents and 40 cents.

When the individual-demand schedule for commodity A is plotted in Fig. 6, it is seen to be a diagonal straight line. With every variation in price, the change in the corresponding amount is exactly proportional. The elasticity of demand is unity throughout. As the price of commodity B falls from 40 cents to 15 cents, the amount demanded increases less than proportionally, as is shown by the fact that the curve falls faster than it moves to the right. The elasticity of demand is less than unity. But from 15 cents to 10 cents it moves to the right faster than it moves downward, representing an elasticity of more than unity.

The collective-demand schedules, and therefore the collective-demand curves, for different commodities are naturally not the same. A *typical* demand curve falls rather rapidly as price declines from the maximum that can be obtained to more moderate levels. Then, as these levels are reached, sales increase rapidly, and the curve tends rapidly to the right. Such demand curves do not exhibit the same elasticity throughout their entire range. Although there is much variation, it may be said that most demand curves are very elastic at high prices (high, of course, for the commodity in question), moderately elastic at moderate prices, and inelastic at low prices. A demand curve is likely to show elasticities of great magnitude in its upper range, pass through a range where the elasticity is slightly more or less than unity, and reach elasticity of much less than unity before buyers refuse to take any more of the commodity.

The principal conditions that affect elasticity of individual demand are as follows: First, the demand of most individuals of the same income class for any given good is usually elastic when the price of that good is high, that is, high in their opinion. When the price of the good is low (in their opinion) the demand is usually inelastic. When price is high they buy sparingly of the good and either go without it or put a substitute in its place. As it falls in price they buy liberally and substitute

it for other goods (which have not changed in price), and when it is low in price they cannot be induced to consume additional amounts unless the price is very much reduced. It may be well to repeat at this point that in modern civilizations practically all commodities have substitutes. Whether an individual regards any good as high in price depends on a very great variety of conditions. Rye bread may be a readily acceptable substitute for wheat bread in certain countries and not in others; in still others rice and other starchy foods may be regarded as more desirable than either.

It has often been said that the demand for a luxury (what an individual or group regards as a luxury) is usually elastic and the demand for a necessity (what an individual or group regards as a necessity) is inelastic. But all that this means is that the demand curves for certain commodities have such a configuration that the individual will buy little of a given commodity, although he would like to consume more of it, because he regards its price as too high. If its price fell he would consume it rather rapidly. This good is a luxury to him. Another good he regards as low in price; he feels that he may consume it freely at the ruling price, and he would not buy very much more of it if its price fell considerably. This good is a necessity to him. To illustrate: most Americans of the lower middle class regard a motor car as a necessity. Nearly all own and use one or more. Obviously they do not believe that cars are high-priced or they would not buy them by the million each year. One often hears the statement that "a car is a necessity." In most European countries people of the same class commonly regard a motor car as a luxury. In most of these countries it is true, of course, that a car of the same quality costs more to own and operate and that the lower middle class has a smaller real income. Yet these same people will often buy opera tickets and good wines, which would commonly be regarded by the corresponding group in America as luxuries.

Second, the demand for any good is likely to become elastic at and above the price where a substitute begins to compete with it. That is, in terms of Fig. 5, if a substitute for the commodity came into the market (curve *IV*) at 40 cents, the curve would lie more to the left and might even intersect the axis *OY* at 60 cents. The demand for a good is likely to become more elastic at and below the price where it becomes a substitute for another good. That is, if the commodity (curve *IV*) became a substitute for another at 20 cents, the curve would fall less rapidly toward *OX*. There is scarcely any substitute for white bread among a majority of the population in the United States, and accordingly the

individual demand for it is relatively inelastic; but the demand for a single brand of breakfast food is elastic by comparison because other brands can take its place, just as it can take their places, in the consumption of most individuals. A great change in the price of one brand is likely to be followed by a greater change in the volume of purchases than would follow from a change in the price of salt, brooms, telephones, and hundreds of other articles or services of daily use for which substitution is difficult.

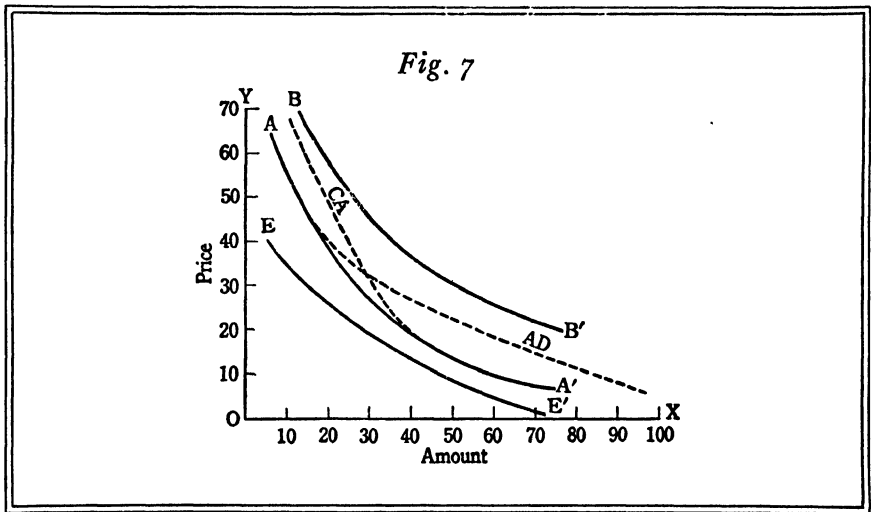
In the third place, the demand for a good tends to be more elastic if it can serve us in more than one capacity. This statement is more generally true for raw materials than it is for finished goods because, after all, the latter as a rule are not adaptable to a large variety of uses. But a good such as coal has a multitude of specific uses in industry in the manufacture of a vast variety of commodities. Fourth, the demand for any commodity is likely to be inelastic at very high and at very low prices. When it is costly, people persist in using small amounts; and when it is cheap, they buy it to the point where the desire for it is nearly satiated. Fifth, demand is likely to be inelastic when consumption is greatly influenced by habit and custom.

The elasticity of the collective-demand schedule is determined by the same conditions that affect the elasticity of its component individual schedules. But it cannot agree with them all. It may be more elastic at a given price than some and less elastic than others. If new buyers enter the market in large numbers at a certain price, the elasticity of the collective schedule may be greater than that of any of its components.

The elasticity of an individual-demand schedule cannot be important, since even a small dealer supplies a number of buyers. It is the elasticity of the collective-demand schedule with which he is concerned. Upon the degree of elasticity of collective demand depend a number of important business decisions. Would it be profitable to cut prices in order to expand and take advantage of the economies of larger production? If the demand is elastic, the prospects of securing enough additional sales to enable the producer to reduce costs are very good; but if it is inelastic, they are poor. Or suppose that the producer's costs are rising. Should he prepare to reduce output greatly as prices are forced up by rising costs? If demand is inelastic, he need not reduce greatly; but if it is elastic, he may be forced to shut down a large part of his plant. The demand for many farm products is inelastic. Hence a change in their supply causes a greater fluctuation in price than would be the case with manufacturers.

CHANGES IN DEMAND

Since demand has been defined as the amount of a certain commodity an individual or a group of individuals stand ready to take at a given price during a given time and in a given market, an increase in demand must mean that the buyers will take a larger amount at a given price; and a decrease in demand, a smaller amount. When the attitude of the buyers toward the good so changes that they will take more or



less at one price in their schedule, it usually happens that a considerable part of the demand schedule and demand curve will be altered. Thus, if we assume that the demand for a good as represented by the collective-demand schedule on page 103 has changed from 140 units at 40 cents to 160 units, we should expect to find that more would be taken at 30 cents and 20 cents, as well as at 40 cents. When it is said that the demand for a good has changed, it is implied that a considerable portion of the demand schedule has moved. But it must not be assumed that all parts of the schedule and of the corresponding curve have changed in the same proportion. In Fig. 7 the original position of the collective-demand schedule is represented by curve AA' . The curve BB' shows an increase of demand throughout the entire schedule, and EE' shows a decrease. Such modifications as CA and AD represent partial shifts in the original schedule.

A common misconception of the meaning of a change in demand is illustrated by the statement, "Because of a decline in price, the demand

for automobiles has considerably increased." As far as the statement goes there has been no increase in demand in the manner in which we have defined it. When we speak of an "increase in demand" we mean that the entire schedule of amounts that consumers will buy at given prices has increased. In terms of the curve AA' in Fig. 7, we mean that it has moved to the position of BB' . But in common usage an increase in demand often means simply that more is purchased because the price has been reduced. Obviously, if the price is at 40 (curve AA') only about 20 units can be sold. But if the price is dropped to 20, then 40 units can be sold. Yet there has been no increase in demand.

An increase in the volume of total sales of any commodity in any market does not always show that an increase of demand has taken place. If the sellers have reduced the price of the commodity at the same time that the increase in sales takes place, we should infer that the additional sales were brought about by increased purchases by old buyers who already stood ready to buy more if the price should fall, and by the entrance of new buyers who previously stood ready to take the commodity if the price fell low enough. The true test of whether demand has actually increased can be made by putting the price back to the old figure. If the volume of sales is found to be greater than before, then we shall conclude that demand has increased. But when sales grow with the reduction of price, we cannot be sure that the additional purchases are not the result of a latent demand that existed before the sellers decided to cut the price.

SUPPLY AND SUPPLY SCHEDULES

In preceding sections the point has been stressed that demand does not mean merely amount purchased and that it should not be taken to mean readiness to purchase a given amount unless the price at which potential buyers are willing to take that amount also is stated. The same points must be borne in mind in analyzing the concept of supply. There are, it is true, instances in which stocks of goods are thrown upon the market to be sold at auction to the highest bidder. In these instances the price at which the goods are offered before the sale opens is any price above zero. Actually there are occasions on which goods are so offered. But the sellers usually have one or more other opportunities to dispose of their stocks if the auction does not call out buyers at higher prices. Hence we conclude that, as in the case of demand, supply invariably implies a relationship between amount and price.

Supply is therefore defined as the amount one seller or a group of sellers stand ready to offer during a given time, in a given market, at a given price. The amount one seller will offer is an individual supply, and the amount all the sellers taken together will offer is a collective supply.

The price at which any seller stands ready at any time to offer a given amount of any good is his *supply price* for that amount. And the price at which the sellers in the market stand ready to offer a given amount is the market, or collective, supply price. The individual-supply schedule is a list of the varying amounts an individual stands ready to offer for sale at different prices. The collective-supply schedule for any commodity is made up, like the collective-demand schedule, of the individual schedules. When supply schedules are represented by supply curves, it is found that instead of slanting downward from left to right the curves slant upward from left to right; that is, as price rises, the tendency is for offers to become more numerous and for more, and not fewer, goods to be put upon the market.

Supply schedules and curves for different sellers of the same good are usually somewhat unlike, just as the demand schedules of different buyers are unlike. Hence the collective-supply schedule is a composite of unlike and divergent individual schedules. The supply schedules, both individual and collective, for different goods are also unlike; that is, a given increase or decrease in the prices of several goods will cause the amounts offered to increase or decline in different proportions. The differences among individual schedules are illustrated in Table 17, in which A, B, and C are amounts of the same good offered by different sellers in the same market.

TABLE 17. Individual-Supply Schedules

PRICE	SUPPLY A	SUPPLY B	SUPPLY C
10	2000	2000	5000
20	2500	2000	5000
30	3000	2000	5000
40	3500	2250	5000
50	4000	2500	5000
60	4500	2500	5000
70	4500	2500	5000

In the case of highly perishable goods already in the consumers' market, the sellers can scarcely be said to have supply schedules. Any amount of the good they happen to have on hand they must either sell

or allow to spoil. If competition is perfect, they will sell the entire stock for whatever price it will bring. This conclusion appears contrary to fact, since it often happens that dealers in perishable goods maintain their established prices and sell whatever *quantity* they can, allowing the remainder to spoil. And it is argued that this course is pursued by the dealers because (1) it is more profitable for them to sell a smaller amount at a higher price than they could obtain for their entire stock, and (2) because they are afraid of "spoiling the market." The first reason requires for its validity an assumption that competition is far from perfect. Wholesale dealers in fresh fruits might indeed find it more profitable to sell three fourths of their supply at a price of five dollars a unit than the entire stock at three dollars a unit. But if every wholesale dealer in a large market had equal access to every retailer, and if every retailer had immediate access to the consumers of several of his competitors, no dealer could sell any of his supply at five dollars, for so long as there existed any unsold stock in the hands of any dealer, he would try to take business away from his competitors by lowering the price a trifle. The competitors would reply by lowering theirs a bit more than his, and the final result would be the fixing of a wholesale price at three dollars. Why, then, do we find that dealers maintain prices and allow goods to spoil? The answer is that each wholesaler has a line of retailers to whom he sells regularly, and each retailer has a group of consumers who customarily patronize his store. If the wholesalers find that they have overbought, they tend to sell what they can at the established price or a little below, because that policy yields them a higher gross return and a higher profit, as will be explained presently. They know that other wholesalers cannot quickly invade their line of retailers and take away the business. The retailer has not established relations with many wholesalers, and he must buy perishable commodities almost daily and in small amounts. For both reasons he tends to take these commodities from the wholesaler at the latter's price, regulating the amount he takes according to the amount he thinks he can sell at the wholesale price plus his customary margin.

The conditions just described, which are very common in business, give rise to what is called "imperfect competition" or "monopolistic competition." Under these conditions the sellers sell in markets that are only partly common to all and partly monopolized by each seller. The prices paid and received are not, therefore, strictly competitive prices. Yet they are not strictly monopoly prices, as will appear in the discussion of such prices in Chapter XII.

A condition in which dealers maintain price cannot force a permanent and wide divergence from competitive price. The growers will bring the commodities to the wholesale market as long as they can obtain a price that will cover costs of harvesting and transportation. The wholesalers will lower their buying price and presently their selling price, and the consumer will buy at a price somewhere near the price that will clear the market.

The phrase "spoiling the market" means that the seller may change the consumers' habits of consumption and their buying behavior to his disadvantage. For some commodities the consumer comes to have in mind a price at which he will buy freely, another price at which he will buy sparingly, and still another at which he will not buy at all. And these prices are much affected by his experience in the past. If the consumer has been accustomed to buy a fresh fruit in season at from 15 cents to 20 cents a quart, and if for a period of a week the price drops to 10 cents, the dealers may have difficulty in selling to the same consumers the following week at 15 cents or 20 cents. Here we encounter a very complex problem. Account must be taken of the income of the typical consumer, the number of substitutes for the particular good in question, the stability of price in the past,—that is, whether the consumer has been accustomed to pay rather stable prices for this good or has been accustomed to pay varying prices,—the extent of competition among retailers, and many other facts.

Some goods that do not deteriorate physically are subject to rapid deterioration of utility with the passage of time. Such "style goods" as articles of dress are typical instances. A retailer who has purchased a stock of goods of this character must sell it before the style changes or the season in which it is wanted has passed. And here again the supply is fixed to a great extent, and the seller must take whatever price he can get. Likewise the dealer will try to protect himself by maintaining prices even to the point of allowing some of the commodity to be held past its season.

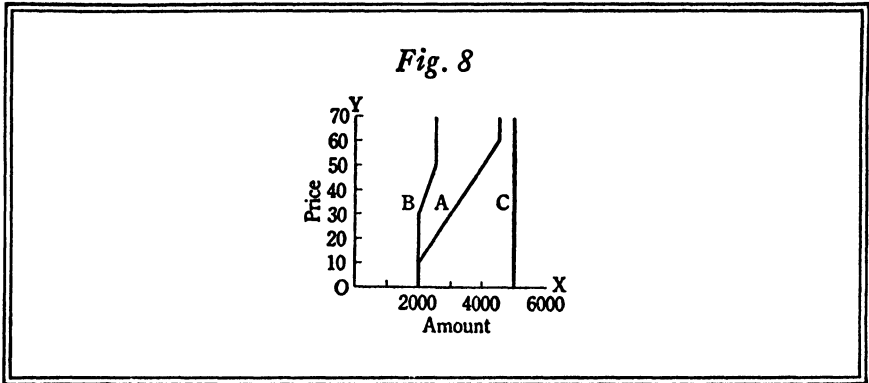
There is of course a tendency for manufacturers and dealers of all sorts to place themselves in a position where they can either withhold goods from the market if the prices are deemed too low or augment stocks if prices are found to be higher than expected. Refrigeration helps to accomplish these ends for perishable goods, and large storage facilities at central marketing points do the same for nonperishable goods of all sorts. For some kinds of manufactures, factory capacity above normal requirements serves the same ends.

For relatively durable goods not affected by changes in style, supply schedules of some range do exist. A farmer, as a rule, must sell his crops sometime during the year following the harvest. But he may distribute his sales within this time according to his judgment of the profitableness of selling at one time rather than another and according to the cost and convenience of marketing crops at one time rather than another. Still another consideration that may affect his decision to sell is his need for funds to pay outstanding debts incurred in producing the crops. Hence he may sell a part of his wheat direct from the thresher, reserving the rest until he thinks the price is the best he is likely to get.

In case the goods are half-manufactured goods or durable finished commodities, supply schedules exist; but they are doubtless much restricted under competitive conditions. For although the goods are not perishable in a physical sense, producers and dealers do not wish to hold them very long in storage because by so doing they incur the risk of a decrease in price and because storage and interest charges accumulate while the goods are held in stock. Hence when confronted with a drop in the demand prices of particular wares, dealers ordinarily sell the stocks on hand for whatever prices they will bring. It is also to be observed that a sudden unanticipated increase in the demand for a given commodity will not always quickly call forth a larger volume of goods. The flow of commodities into the market is regulated in advance by the producers, and time is required to make additions to the rate at which they are fed into the market.

There are certain goods and services for which the supply schedules are very extensive; that is, they run through a wide range of prices and amounts. Dwelling houses, used automobiles, furniture, and other goods frequently offered for sale by consumers who have the alternative of selling or using them are commonly offered at greatly varying prices. Because different people are certain to attach to these goods different values for personal use, there is less uniformity than in a market where only dealers are sellers. The supply schedule for labor also is extensive. Here the wage rate is weighed against the irksomeness of labor. At low wages (low relative to the established rate) many laborers refuse to work at all and others refuse to work many days in the month, whereas higher wages call forth more workers and induce many men to work overtime and more days in the month.

In Fig. 8 are plotted three schedules of supply shown in Table 17 on page 111. The perpendicular lines extending from the termini of



curves *A* and *B* serve only to show that after certain prices have been reached, amount ceases to vary with price. In the case of supply *C*, the amount is fixed and there is no covariation and no curve, but simply an ordinate—that is, a perpendicular line extending indefinitely above the axis *OX*.

THE DETERMINATION OF SUPPLY

In our discussion of the types of supply schedules, we noted that sometimes a stock of goods is thrown upon the market to bring whatever the buyers will give for it. This is sometimes called supply without reservation, because no part of the stock is withheld if the price is unprofitably low. In many other cases sellers do reserve goods; that is, they hold them at certain prices below which they refuse to sell. The conditions that govern the fixing of dealers' supply prices are very complex, and we cannot go into them at length at this point. Briefly stated, the following are the most important: (1) dealers' estimates of the prices consumers will probably pay for the amounts available now and at a near future date; (2) dealers' estimates of the amounts of the good that will come upon the market in the near future; (3) the prices of substitute goods; (4) present and probable future costs of selling and storage; (5) the prices at which competitors now offer, or are likely to offer, the good in the near future; (6) present and future estimated cost of buying goods from other dealers or producers.

The seller in a modern competitive market must be prepared to change the reservation prices he has set upon his stocks; for such reservation prices are largely estimates, on the part of the seller, of the prices the market will pay for the goods. But no one can predict these prices accurately in advance of actual experience. Neither the exact

demand prices of the buyers nor the number of persons who will buy can be forecast, and the amounts of goods in the hands of manufacturers, competing dealers, and other marketing agencies cannot be ascertained easily or accurately. The supplies of substitutes and the degree to which substitution can take place also are indeterminate. The practical method of pricing seems to be that of "cut-and-try." If the price first tried is too high, few goods will move; and that is sufficient warning to the sellers to reduce prices. It often happens that they must reduce them so low that the goods they have stocked become occasions of loss. If this is the case, smaller quantities will be stocked in the future; and in the long run, the goods either will sell for enough to enable manufacturers and dealers to make and handle them profitably or will disappear from the market.

The difficulty of ascertaining in advance the proper reservation prices causes the supply schedules of many goods to be extremely variable; that is, the schedules move or change position very frequently. There may be no such thing as a stable supply price. Of other goods, particularly branded goods and smallwares, reservation prices change only infrequently, the manufacturer preferring to regulate the supply so as to feed the market at a rate that will maintain the established price. This policy has recently come to be known as that of "administered prices." The prices of electric current for domestic use, of telephone service, of steel rails, and of numerous other goods often remain fixed for many months. In all these cases, as well as in the case of branded goods, imperfections of competition, which may amount to a considerable degree of monopolization, account, in part, for price stability. It must not be inferred, however, that monopoly prices when intelligently adjusted are necessarily stable. This question will be dealt with later, under the discussion of monopoly price.

CHANGES IN SUPPLY

Increase in supply means that more goods are offered at a given price than before the increase took place, whereas a decrease in supply means that fewer goods are offered at the given price. An increase or a decrease of supply at one price in the supply schedule usually means that the supply has increased or decreased at adjacent prices. Hence an increase of supply is illustrated graphically by placing the supply curve farther to the right and below its original position, a decrease by placing it farther to the left and above its original position. But

it must be remembered that because the supply curve (or schedule) has changed for certain prices, it does not always follow that it has changed for all prices.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Suppose that you found that the price of gasoline was 15 cents in a certain city and 20 cents at a roadside filling station 50 miles distant in the country. Enumerate the differences in demand and supply that might account for this variation in price.

2. Criticize the statement that governmental restriction on the production of a crop is unwise because "it is contrary to the law of demand and supply." Is government the only agency that arbitrarily influences supply?

3. Is the statement that "demand and supply" determine price an economic law? Is it a generalization?

4. If you were the only producer of a good, which would be of more significance to you: (1) the absolute amount by which price would fall (or rise) as the quantity increased (or decreased), or (2) the relative amount (elasticity) by which it rose or fell?

5. For what reasons might the demand curve of one individual be different from that of another individual?

6. Make a list of the commodities for which your own demand at the current prices is (1) elastic, (2) inelastic. Can you explain the difference in practical terms?

7. Make two lists of the goods you frequently buy, one including those that fluctuate frequently in price, and the other those that are relatively stable.

CHAPTER VIII · The Price-Making Process: Market Price



RESERVED AND UNRESERVED SUPPLY

The supply of a good is said to be *reserved* if the sellers, before entering the market, set prices upon their respective offerings below which they will not sell. The prices at which the sellers reserve their offerings are called *reservation prices*. In order that sellers may reserve their supplies and set reservation prices upon them, they must be able to hold back their stocks if prices are, in their judgment, too low, and to push them forward when prices are satisfactory. Reservation implies, therefore, the existence of storage facilities and waiting power on the part of the sellers, or other uses for the good.

The supply of a good is *unreserved* if the sellers throw it upon the market without setting any price whatever upon it. Often reservation prices are not made public. The sellers simply ask the prospective buyers, "What will you give for this commodity?" But in a majority of these cases those who offer the commodity have privately determined that if the prices do not meet their expectations, they will reserve their stocks either until the prices rise to what they are willing to take, or until their reservation price has been proved too high.

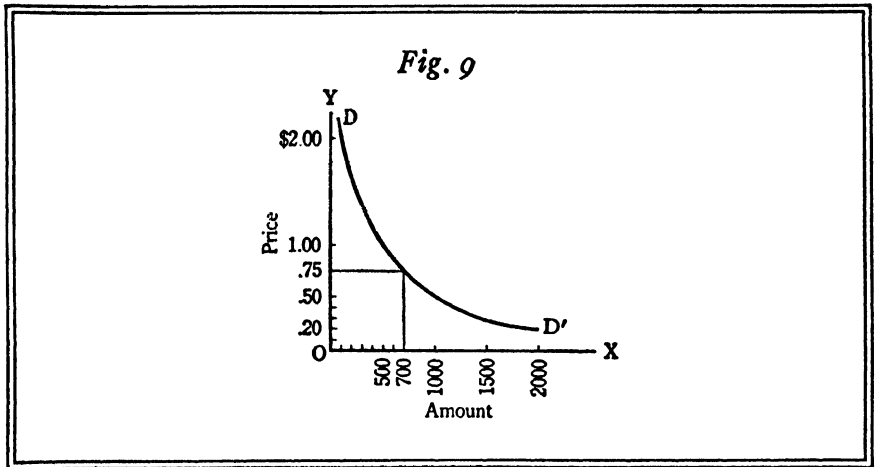
PRICE WITHOUT RESERVATION OF SUPPLY

The determination of price without reservation is the simplest price phenomenon encountered in the market, and in many markets it is nearer the actual conditions than is the assumption of reservation of supply. Hence the making of price without reservation will be explained before the more complicated problem of price with reservation of supply.

The conditions of price-making without reservation are as follows: (1) a competitive demand for the good is given, and (2) varying quantities of the good of which the price is being studied are thrown onto the market from day to day by a large number of independent sellers to sell for whatever they will fetch, any price above zero being acceptable.

The statement that a large number of independent sellers offer the good means that there is a high degree of competition among them.

The effect of throwing an unreserved supply onto the market depends upon whether the buyers have more or less complete information of the quantity actually available. If they believe that the quantity is smaller than it actually is, they may pay high prices for a time and, when the amount of the good available becomes known, lower their bids for it. It may happen also that they have overestimated the supply and



that the initial low price will be followed by higher prices later. But when this preliminary higgling has revealed to all buyers the actual amounts for sale, price will settle down to a more stable level. It is with the stable level that we are concerned.

If the demand for a good is given, and if a definite amount is offered for whatever it will bring, then the level toward which the higgling will tend to drive price is determined by the collective-demand schedule for that market and by the amount offered.

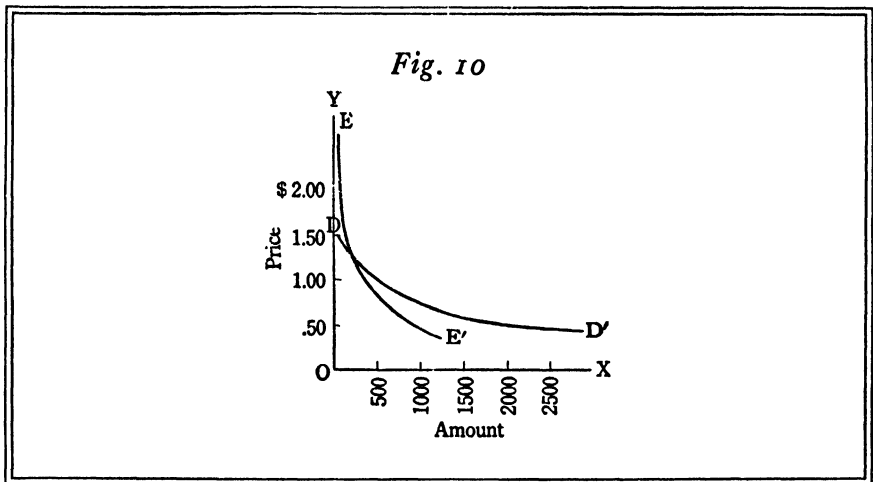
In Fig. 9 the demand schedule, which is represented by DD' , varies from 100 units at \$2 to 2000 units at 20 cents. If the sellers throw 500 units on the market without setting a reservation price, the good must sell for \$1 a unit after the price has become stable, and the gross receipts from that amount will be \$500. If 1000 units are offered, the price will be 50 cents a unit and the gross receipts \$500, or exactly the same as though 500 units had been sold. And if 2000 units are offered, the price will be 20 cents and the gross only \$400, or \$100 less than if one fourth that quantity had been marketed.

It is evident that when the demand curve is fixed,—that is, does not move to the right or the left in terms of the diagram—any increase in the amount offered will lower the price, and any decrease will raise the price. These conclusions follow from the fact that the demand curve for practically all goods slants outward from *OY* and downward toward *OX*. In the illustration it appears that a large amount not only brings a smaller price per unit, but also a smaller gross return. But this is not always the case. In the first place, the smallest amount that can be sold with demand as represented in Fig. 9, 100 units, will sell for \$2 a unit and yield a gross return of only \$200, which is less than if 500 units had been offered. A more careful inspection of the prices at which various amounts can be sold shows that the maximum gross is received by the sellers when approximately 700 units are sold for a price of 75 cents, which will return \$525 gross. Any amount in excess of 700 units yields less in total money return, and any amount below 700 will do the same. The reason for this condition is plain. At the upper end of the curve, price is very high, but only a negligible quantity can be sold; at the lower end, large amounts can be sold, but the price is very low. Hence it follows that between the two extremes, at some mid-point in both price and amount, the gross receipts will be at a maximum.

This conclusion applies to all goods sold in markets where the demand schedule has definite limits. If any good were an absolute necessity of life, the demand schedule for it would show that indefinitely high prices would be paid for small amounts, since life itself would depend upon it. The demand curve, instead of meeting *OY*, would become nearly parallel to it and extend upward indefinitely. Under this condition a relatively small amount would bring a fabulously high price and yield the maximum gross receipts. This condition of demand is shown by the curve *EE'* in Fig. 10.

The opposite situation to the one just described is encountered if it is assumed that at low prices the demand curve flattens out and becomes nearly parallel to *OX*. Curve *DD'* in Fig. 10 shows that the larger the amount put upon the market, the larger will be the gross receipts. Even though the price falls with each increase in sales, it falls slowly, while amount increases rapidly. Of course there is no commodity the demand for which is so imperative that after a certain point has been reached amounts taken increase very greatly and indefinitely with each minute decline in price. All demand schedules reach zero price much short of the point where more than a small proportion of the nation's purchasing power is used to buy any single good.

But it is not contrary to fact to say that at certain points in the demand schedules for some commodities the amounts buyers stand ready to take increase so rapidly with small changes in price that gross receipts also increase rapidly. If the demand for a good is sufficiently elastic between two price points, gross receipts must increase. Such a condition of elasticity depends on the possibility of direct substitution of the good for one or more other goods by a large number of people. But when this condition applies to a certain good, it must be true, not

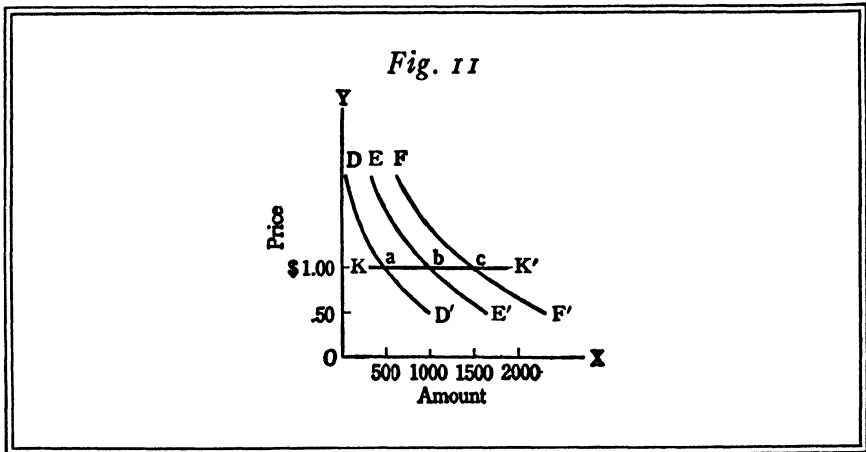


only that the good is readily accepted by the consumer in place of other goods, but also that the producers of the other goods cannot reduce their prices enough to prevent the substitution.

The effect of high elasticity upon price and gross receipts must be sharply distinguished in this analysis from the effect of an increase of demand. In Fig. 11 the demand schedule moves from an original position DD' to EE' and FF' , the latter two positions both representing increases in demand at all points on the demand schedule. When the demand is at DD' , 500 units sell for \$1 and gross receipts are \$500; when it is at EE' , 1000 units sell for \$1 and receipts are \$1000; and when it is at FF' , 1500 are sold and receipts are \$1500. The conditions represented in Fig. 11 are quite different from those illustrated in Figs. 9 and 10. In the latter two figures the demand curve was assumed to remain fixed while the amount offered increased or decreased. In Fig. 11 both the demand and the amount offered have changed, and as a result of this double change, the price has remained the same. But the stability of price is due solely to the fact that both the demand and

the amount offered have changed. If the amount had lagged behind demand, price would have risen; if it had outstripped demand, price would have fallen.

Such increases in demand as are presented in Fig. 11 result from the natural growth of population, from immigration, from an increase in the national income, or from changes in habits of consumption involving a reduction in the use of some commodities and the adoption of others. But these are all changes which do not happen within a short range of



time. In a short period of time the sellers do not have the choice of selling either 500 or 1000 units at \$1 a unit. Having produced either quantity and being unable to reserve supply, they must sell for the price determined by the collective-demand curve.

Thus far in the discussion we have assumed that goods are thrown upon the market without reservation, without any price set upon them by the sellers. It may well be objected that sellers practically never do this. Although the sellers usually set up, either openly or in their own minds, the minimum prices they are ready to accept for different amounts of the goods they sell, several conditions actually encountered in economic life detract somewhat from the significance of reservation prices. In the first place, the sellers may be unable to forecast the demand. They may expect it will be sufficient to carry off 1000 units at \$1 and produce \$1000 gross; but when this amount is brought to market, it may be found that instead of being at *EE'* (Fig. 11), demand is at *DD'*, and the stock that has been produced must be sold for a gross of \$500. If this is the case, it is futile for the sellers to hold the good at \$1, provided it is necessary to dispose of the entire stock within the period

of time to which the demand curve DD' is applicable, and provided perfect competition exists. But if the good is durable and if future production can be controlled, the sellers may find it more profitable to offer something between 500 and 1000 units and hold the remainder for a higher price in the future. If the future supply cannot be controlled or if the stock is perishable, they must take 50 cents. To be sure, reservation prices will be attached to the good in either case, but in the latter they are derived entirely from the demand as shown by the prices at which consumers are willing to buy. For example, in the past farmers have not been able to store more than a small portion of their crops. If the wheat crop happened to be unusually large in any given year, it had to be sold for a price determined by the world demand for wheat and by the amount of the harvests in all the wheat-producing areas of the world. A small amount was held in storage; but since the next year's planting could not be reduced materially, and since next year's yields, which depended largely on the weather, could not be effectively controlled, this procedure was not of material benefit.

The danger from excessive production is not confined to agriculture. Within a short period of time the market for many commodities is limited. Consumers' demands have not a high degree of variability over short periods of time, because, among other reasons, consumption is so greatly controlled by habit and custom. Hence the principles of price applicable to supply without reservation are significant for price analysis in all types of production.

THE SIGNIFICANCE OF SUPPLY PRICES

The system of uniform prices for buyers, which has now largely superseded the older system of individual bargaining throughout retail trade in the United States, is essentially a device for economy and for retaining good will. It saves time formerly wasted in making individual bargains; and it pleases the buyer, who likes to believe that he is receiving equal treatment with all the other patrons of the store. The buyer must pay the merchant's price or go without the good. This system gives the appearance of greater power to control price than the seller actually possesses, for it must be remembered that a seller cannot maintain a predetermined price unless he can also limit the amount the entire market offers for sale. If large numbers of the consumers do not want to pay the price asked, they usually can force the seller to reduce it by refusing to buy.

The actual control over prices, which the seller exercises quite as much as the buyer, arises from the control of quantities and not from the power to set reservation prices. Hence, if any producer wishes to control price, he must limit the amount of goods offered for sale. Moreover, if the control is to be accurate and effective, the seller must know in advance the demand at various prices.

MARKET PRICE WITH RESERVATION OF SUPPLY

In the preceding chapter the nature of supply schedules and supply curves was briefly stated. The reasons why different sellers offer, and can economically afford to offer, a good at different prices, even in the same market, are numerous and often difficult to ascertain. They are affected by variation in the location of plants, in wages, in taxes, in climatic conditions, and in local demand. Much of the discussion in succeeding chapters is concerned with these variations. In this simplified discussion of price-making it is merely assumed that supply curves exist. They are not explained.

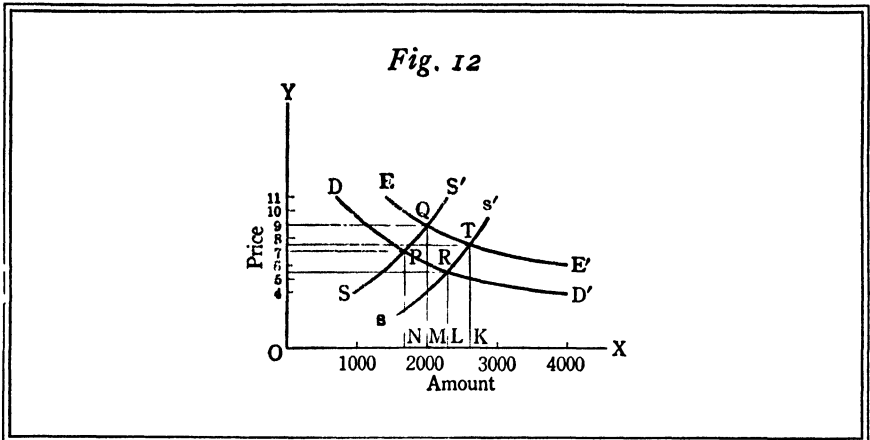
The supply schedule represents the amounts that sellers stand ready to sell at varying prices, and the demand schedule the amounts that buyers stand ready to buy at the same prices. When amounts offered for sale in any market are regulated by reservation of supply, the price at which the good sells depends upon the relation between the demand schedule and the supply schedule. In Table 18 the schedules of demand and supply show how market price is determined under these conditions.

TABLE 18

PRICE (CENTS PER UNIT)	DEMAND SCHEDULE (UNITS)	SUPPLY SCHEDULE (UNITS)
11	700	2300
10	900	2150
9	1100	2000
8	1400	1850
7	1700	1700
6	2000	1450
5	2600	1250
4	4000	950

At the relatively high price of 11 cents the rate of flow of goods to the market (amount offered per unit of time) is more rapid than the rate at which consumers are willing to buy, and at the relatively low price of 5 cents the amount the consumers stand ready to take far ex-

ceeds the amount the sellers care to offer. It follows that at a price of 11 cents the market will be flooded with unsold goods, a condition that sellers cannot endure, while at a price of 5 cents many willing buyers will be unable to make purchases. But at the price of 7 cents the amount sellers offer and the amount consumers stand ready to buy (1700 units) are exactly equal. This is the only price that will bring about an equilibrium between demand and supply. There is, then, in any given



market, under competitive conditions, a single price at which demand and supply are equal; if any other price happens to be set at any time, it cannot be maintained. Price will move to the point where demand and supply are equal.

EFFECTS OF CHANGES IN DEMAND; SUPPLY CONSTANT

In Fig. 12 the demand and supply schedules of Table 18 are represented by the curves DD' and SS' . The equilibrium price is seen to be 7 cents, or NP , and the amount sold, 1700 units, or ON units. When demand increases throughout the entire schedule, the new equilibrium (the supply schedule remaining the same) is at the intersection of SS' with EE' , or at Q . The new price is MQ , or 9 cents a unit, and the amount sold is OM , or about 2000 units. The result is different from the effect of an increase of demand when supply is a fixed stock. In the present illustration an increase of demand with only 1700 units for sale will send the price to about 10 cents a unit. But since the sellers are willing to offer 2150 units for 10 cents a unit before the rise in demand takes place, the price will not rise to 10 cents, but only to 9 cents. It

can readily be seen from the diagram that if demand declines below DD' , the price will fall. But it will fall less if the sellers are in a position to reserve their stocks, as indicated by SS' , than if the amount offered for sale is not reserved.

EFFECTS OF CHANGES IN SUPPLY; DEMAND CONSTANT

An increase in supply is represented by ss' . If demand remains DD' when this change in supply takes place, price will decline to LR , or $5\frac{1}{2}$ cents. The amount sold will increase by the amount NL . Again it is self-evident that if supply decreases—that is, if SS' moves to a position above and to the left—the price will rise above NP , and fewer than ON units will be sold.

EFFECTS OF CHANGES IN BOTH DEMAND AND SUPPLY

If demand and supply both change, as represented by the movement of DD' and SS' to positions EE' and ss' , the new price is KT , or about $7\frac{1}{2}$ cents, and the amount sold is OK , or 2600 units. When both demand and supply are changing, the effect upon price depends on the extent to which they change with reference to each other. If demand increases more than supply, price will rise; and if supply increases more than demand, price will fall. These are the results that follow in well-organized, competitive markets. When competition is imperfect, great changefulness of demand and supply leads to chaotic movements of price. It is possible for an increase in supply that is much greater than the increase in demand to be accompanied in portions of the market by a rise instead of a fall in price, or at least a decline in price that does not result in a cleared market. When this happens, the period of high price (high relative to actual conditions in the market) is followed by a compensating decline in price below the equilibrium point. But such conditions are self-correcting. The sellers finally discover that they cannot dispose of their stocks when price is too high and that they cannot supply the market when it is too low. Hence price moves more or less rapidly in the direction of the equilibrium point.

MARGINAL BUYERS AND SELLERS

In the price-making process certain buyers and certain sellers are likely to be in critical positions with respect to their influence upon total supply and total demand. In the following table, four purchasers are

ready to bid different prices for one unit of a commodity, and four sellers, having one unit each, are ready to offer the good at different prices.

A will buy 1 unit at \$5.00
 B will buy 1 unit at \$4.00
 C will buy 1 unit at \$3.00
 D will buy 1 unit at \$2.00

E will sell 1 unit at \$2.00
 F will sell 1 unit at \$2.50
 G will sell 1 unit at \$3.00
 H will sell 1 unit at \$4.00

Only three units of the good will be sold, and the price will be \$3, for at that price the amount the buyers stand ready to take exactly equals the amount the sellers stand ready to offer. The lowest bid of \$2, by D, and the highest offer of \$4, by H, cannot be satisfied. It will require a price of \$4 to bring H into the market, and at that figure four units will be offered and only two bought; E, F, and G will cut prices, because they stand ready to sell at less than \$4; and when price reaches \$3, buyer C will come in, but seller H will have dropped out. Likewise it can be shown that if the bidding is started at \$2, competition among the buyers will quickly eliminate buyer D. In this illustration C is the marginal buyer; that is, he is the buyer who is most sensitive to a rise in price above the equilibrium point of \$3. He is at the margin, or limit, of the demand that can be satisfied in this market, and any rise in price will cause him to drop out. In like manner G is the marginal seller, the most sensitive of all sellers to a fall in price. He is at the margin of the supply, and any decline in price would cause him to withdraw.

It must not be supposed that either C or G determines the market price by setting his bid or offer at \$3. If either A or E withdrew from the market, a new price would be established, just as it would if the marginal buyer and the marginal seller withdrew. But neither A nor E is likely to withdraw, because the former is getting for \$3 a good for which he was prepared to pay \$5, and the latter is selling for \$3 when he would have taken \$2 rather than not sell at all.

When the total demand for a good is made up of many bids, and when the total supply is composed of many offers, it is the marginal increments of demand and supply upon which attention must be centered in an attempt to forecast price changes. In Fig. 12 the marginal increment of supply is the amount that is put upon the market with a reservation price of exactly 7 cents. This is the portion of the total supply that would be withdrawn if the price decreased below 7 cents. The marginal increment of demand is likewise the portion of the total amount demanded that buyers would withdraw if the price increased slightly above 7 cents.

The marginal supply price of any commodity is the price at which the marginal increments of supply are put on the market, and, at the same time, the price necessary to call forth these final increments. The marginal demand price is the price at which buyers stand ready to buy the final increments actually purchased. It is the price which the buyers think just worth paying for the final increments purchased. In a perfect market the marginal supply price and the marginal demand price exactly coincide. If the marginal increment of demand proceeds from a single buyer, then the buyer is the marginal buyer. But many buyers may participate in the marginal increment of demand, and all of them may reduce their purchases slightly if the price rises. Similarly, the marginal increment of supply may be furnished by a single seller or by several sellers.

EQUILIBRIUM PRICES AND ACTUAL PRICES

In the preceding sections the theory of market price under complete competition has been explained. In this explanation, certain assumptions have been made concerning the conditions under which the selling and buying take place. All these assumptions are realized only if the most highly perfected marketing organization prevails; but if the organization is not perfect, some of these assumptions will fail, and consequently actual price will vary from the equilibrium price. These assumptions must now be examined.

EQUILIBRIUM PRICES UNDER CONDITIONS OF MONOPOLY

In the preceding sections of this chapter, the assumption has been made that there are many rival buyers and sellers. Competitive price will prevail when there are a large number of independent buyers and sellers in any market and when competition is perfect. When, however, there is only one seller of any commodity a monopoly exists. When the number of sellers is few they may agree to limit supply and thus create quite as effectual a monopoly as though only one seller controlled supply. Such agreements are not uncommon, but since they have usually been illegal in the United States, they have been as well concealed as possible. A much more common imperfection of competition is found when one seller possesses some advantage over other sellers in the same market that enables him to restrict output sufficiently to obtain unusual profits but does not give him complete control over supply.

Thus a railroad may have the power to charge higher freight rates than would be obtained if there were sufficient traffic to warrant the building of a competing line. In some small isolated villages there may be only one dealer in a certain type of merchandise, and that dealer may exercise some monopoly power, but his customers can always travel to other villages or towns if the difference in price is worth the trouble, and hence his control of supply is not complete. These latter illustrations pertain to what is called imperfect competition or monopolistic competition. A simple monopoly exists only when one seller or group of sellers has complete control of the supply of a given commodity. The analysis of price under conditions of simple monopoly is the subject of this section.

A monopolist cannot control price at will. All that he can do under most conditions is to control supply. Under certain conditions, he may influence demand through advertising or by some other method of breaking down sales resistance. But he cannot control demand as he can the amount of the good offered. Here it is assumed that the demand curve for the monopolized product is exactly the same as though perfect competition prevailed. If the monopolist takes full advantage of his control of supply, he will set a supply price that will yield the maximum monopoly revenue or monopoly profit. If the monopolist knows the position of the demand curve, his supply curve will be like curve *C* in Fig. 8, page 115. That is, the monopolist will offer exactly the quantity that will yield the maximum monopoly profit. Since we have not introduced the concept of cost in this chapter, it will be convenient to assume that the monopolist has not incurred any cost of production, as he would not if he owned and leased the only mines from which a rare mineral was produced or if he owned the patent for a highly valuable chemical process, the employment of which he leased to others. In later chapters, where the relation of cost to price is considered, price-making under the influence of both cost and monopoly will be analyzed.

The question to be answered at this stage of the discussion is, How should the monopolist determine the amount that will yield the maximum monopoly profit? To answer this question exactly we must refer to Chapter VII, in which the characteristics of the collective-demand curve were described. On pages 106–107 it was found that, in general, demand is elastic at high prices and inelastic at low prices. It was pointed out also that before unit elasticity is reached, as the demand curve descends, the product of price times quantity increases, and that this product reaches a maximum when elasticity is unity and

thereafter declines. The reason for this can be illustrated by a demand schedule and a diagram.

From Table 19 it can be seen that the total revenue increases as the price falls until a maximum of \$25 is reached and thereafter declines until at zero price no revenue is produced. In Fig. 13, a diagram similar to those previously employed, DD' represents the collective-demand schedule. Since it is a straight line, it shows very high elasticity near OY and very low elasticity near OX . The elasticity of DD' at the mid-point P is unity. The curve Dd , called the *marginal revenue curve*, shows the revenue added by each additional item sold. According to Table 19 the first item adds a revenue of \$9 because by the sale of that item the total revenue increases from zero to \$9. The second item adds the difference between the total derived from the sale of 1 unit and the total derived from the sale of 2 units, or \$7, and so on. So long as the marginal revenue is positive—that is, so long as added sales add to the total revenue—the monopolist should increase his output. But when the marginal revenue falls to zero he should cease to expand sales. If he should sell more—say 6 units—the marginal revenue would be -1 and the total revenue \$24 instead of \$25.

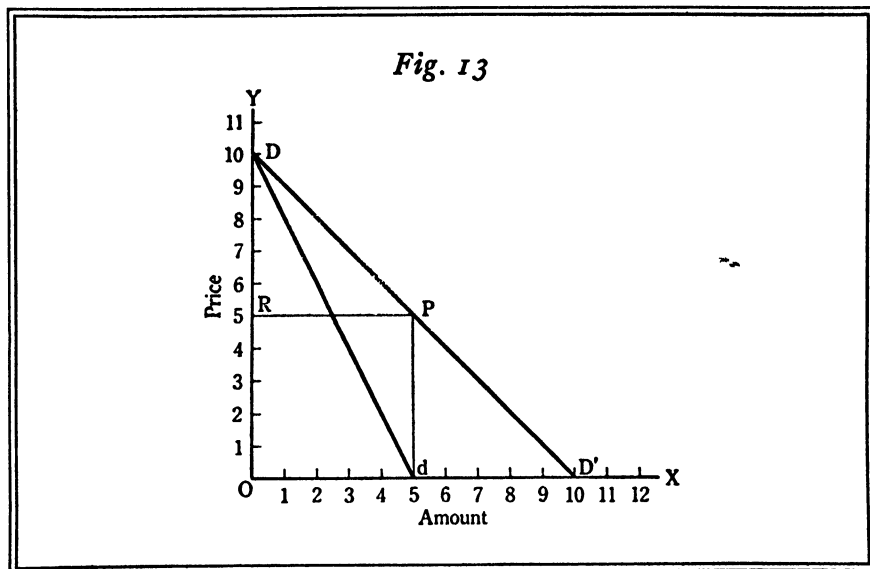
TABLE 19

QUANTITY SOLD	PRICE (DEMAND)	TOTAL REVENUE (PRICE × QUANTITY)	ADDED OR MARGINAL REVENUE
0	\$10	\$0	\$0
1	9	9	+9
2	8	16	+7
3	7	21	+5
4	6	24	+3
5	5	25	+1
6	4	24	-1
7	3	21	-3
8	2	16	-5
9	1	9	-7
10	0	0	-9

In Fig. 13 the relationships between the collective demand, marginal revenue, and monopoly profit are more accurately represented. Certain discrepancies between the table and the figure are easily noted. The table shows marginal revenues of 9, 7, 5, and 3 when 1, 2, 3, and 4 units are sold, while the figure shows marginal revenues of 8, 6, 4, and 2, respectively. This discrepancy arises from the fact that we are dealing with infinitesimal quantities, which cannot be represented accurately

by numbers varying as much as one unit (as in Table 19). The curve Dd in Fig. 13 is accurate, and the table is not, because the curve does correctly represent infinitesimal increments. The table is useful, however, because it shows roughly how the marginal revenue is computed.

The difference between the supply price of a monopolist and of a large number of competitors can now be explained by means of Fig. 13.



Suppose, first, that several competitors have for sale a stock of 8 units and that they have neither the inclination nor the power to set a reservation price on this stock. That is, there is no supply curve. The stock must be sold for \$2 a unit, and the gross will be \$16. Now assume that a monopolist holds this stock for sale and that aside from his monopoly power he too has no reservation price. But because he has control of the supply he will offer only 5 units for sale at a price of \$5 a unit, yielding a gross of \$25. The amount he will sell is indicated by the point d , where the marginal-revenue curve Dd cuts OX . He will sell neither more than Od units, because any added amount causes the added revenue to become a minus quantity, nor less than Od , because until that amount is reached each added unit adds to the gross revenue. The price will be dP and the gross $RPdO$, the maximum obtainable.

SUMMARY ON MARKET PRICE

In this chapter, the equilibrium of demand and supply under conditions of perfect multiple-supply competition and simple monopoly have been explained. The reader is doubtless aware that perfect competition is seldom found in the real world. It is likewise true that conditions that permit one seller to control supply without fear of intrusion by outsiders also are rare. There is, in most retail markets, a degree of monopoly arising out of personal relationships between buyer and seller, location of stores, special services, advertising, and brands. All these conditions enable a given seller to hold a portion of his customers even when competitors offer the same or equivalent goods at somewhat lower prices. A seller is said to have "his customers" in much the same sense that a physician has a clientele of those who consult him without giving much thought to the fees charged or services rendered by others in the same profession.

For the most part the prices of goods, however, are chiefly competitive. There are hundreds of wares that cannot be sold at different prices by different dealers unless there are also differences in the quality of the goods or in the services rendered by the sellers.

In the markets where dealers buy their stocks or manufacturers buy equipment, conditions are even more competitive than in retail markets. Public-utility franchises, patents and copyrights, and seductive advertising may give rise to fairly complete monopolies, but even in these instances the fear of competition by the intrusion of new sellers is nearly always present. The railroads and street railways have found that exclusive franchises do not shut out the competition of the automobile and the airplane. Patent monopolies are limited in duration, and imitation is difficult to prevent. Even a monopoly based on the fact that a large amount of capital is necessary to establish an economical plant cannot be defended indefinitely, as the example of competition between producers of low-priced motor cars well illustrates. The actual cases of limited competition are not usually cases of complete monopoly. They are rather cases of temporary and evanescent restriction on perfect competition. First one producer or good and then another catches the fancy of the buyer. The result is that first one concern or product enjoys a favored position and then another. Hence, in a world where friction, change, and whim prevail, equilibrium always tends to be established but is attained only at intervals, or perhaps never.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Criticize the following statement: The economist's theory of market price applies to the highly organized markets, such as the produce exchanges. It does not assist in understanding price in retail markets.

2. A branded good that is widely advertised sells for exactly the same price in all stores throughout the country. Is this a competitive price? Is it an equilibrium price? What would probably happen if the volume of sales was too small to take the output of the manufacturer?

3. Criticize the following statements. Which are correct and which incorrect?

a. If price rises, demand falls off.

b. Demand depends on price.

c. Price is partly determined by competition among sellers.

d. Price is partly determined by competition among buyers.

e. When there are many buyers and only one seller, price is not determined by demand and supply.

4. The extent of the change in price that will follow from a change in supply depends upon the slope of the demand curve. Why?

5. Are the conditions or events likely to bring about a change in the demand for wheat, rayon, and automobiles the same? How different?

6. Are reservation prices attached to the annual wheat crop by farmers? to men's straw hats by retailers? to old paintings by dealers? to electric current?

7. In villages we often find that there is only one dealer in each line of retailing. Yet the prices he charges are substantially the same as those charged by other sellers in neighboring towns. Why does the dealer not charge a price considerably above the competitive price? Is the supply that he furnishes the only supply in the market?

8. Do you think that the annual wheat crop of the United States sells at a price that yields the maximum gross receipts to the farmers? Why, or why not? If not, farmers would gain by destroying a portion. Why does such destruction practically never occur unless by government intervention?

CHAPTER IX · Consumers' Demand and Utility



THE ANALYSIS OF INDIVIDUAL AND COLLECTIVE DEMAND

In the preceding chapter the demand schedules of the individual and of the entire market were employed in the explanation of market price. The statement that the more of a good there is to be sold the lower will be the price that people will pay for it was relied upon to explain the determination of market prices. But why individual and collective demands display this tendency was not explained. Moreover, no explanation was offered for the fact that the prices consumers stand ready to pay for some goods are high while those they will pay for others are low. The explanation of market price in terms of demand and supply remains valid even though we should be unable to discover and describe the forces that determine the demand of consumers for finished goods. Observation has amply demonstrated the important characteristics of the collective-demand curve, such as its downward and outward slope, its tendency to vary, and the varying elasticity of demand for different goods.

In all discussion of the forces behind consumers' demand the economist must proceed with caution. At no point is a final answer to the problems of demand which we have enumerated in the preceding paragraph in sight. For reasons that are obvious, the conduct of buyers in the consumers' market is as yet only partly understood. Some of the statements to be made in this chapter therefore are hypothetical. But these hypotheses are not wholly unverified, and they are the most satisfactory explanations of the facts that thus far have been advanced. At present the verification of these hypotheses depends upon three sorts of data: (1) We can rely, to some extent at least, upon the statements of those who are engaged in selling goods to consumers. (2) We can bring to bear upon the problem the findings of other sciences that deal with human behavior. (3) We can rely to some extent upon general observations that any person can verify for himself by noting the actions of consumers. In testing the conclusions of economics about consumers' demand three questions therefore should be kept constantly

in mind: (1) Do these conclusions agree with the statements of those who deal with consumers in the market? (2) Are they in accord with the accepted findings of psychology and sociology? (3) Are they in accord with everyday experience?

It has been said that in studying the demand for consumers' goods we are dealing with human behavior and with human motives. It might seem, therefore, that the subject is primarily one that really lies within the field of psychology and not within the field of economics. But this is not necessarily the case, for economics and other sciences besides psychology are concerned with human behavior. The historian, the lawyer, and the criminologist cannot refuse to consider it even if they would. The investigation of human activities and their causes cannot, therefore, be supposed to be the exclusive domain of psychology. The economist, of course, must accept the findings of psychology as to man's fundamental traits and his fundamental modes of behavior. But not all these findings are relevant to a study of economics. Economics studies only a portion of man's activities, only a limited division of the entire field of human behavior and human motives. It professes to study only those activities having to do directly and indirectly with buying and selling goods and services. As far as the subject of the present chapter is concerned, we shall deal primarily with the exchange of goods for money. Yet it should be noted at the outset that the student of economics is aware that it is impossible to separate economic conduct from other activities of the individual. No individual is entirely an economic man any more than he is a political man. But because we are concerned with buying and selling, we must arbitrarily separate, as far as we can, man's economic conduct from his political, religious, and other activities. Such a separation is an abstraction, but it is a type of abstraction that is common in other sciences. Clearly, then, it is no criticism of our method to point out that under the stress of religious motives men sometimes give away their goods and devote their lives to the service of other people without compensation. We are not engaged in a study of men acting under the dominance of religious or philanthropic motives. But since men are moved by considerations other than economic, the findings of economics must, when practical activities are contemplated, be checked by the conclusions of those who are engaged in the study of other types of human activity. Economics undertakes to explain only economic behavior.

In studying the behavior of man in earning his living the economist is justified in making certain assumptions that are typically true for

the business man, the wage-earner, the investor, and the consumer in their economic activities. We assume that the consumer always seeks to secure the maximum gain for the least cost, whether that cost is in terms of goods or in terms of the sacrifices of labor and saving. We assume that in his economic activities the consumer, the seller, the wage-earner, or the investor will act according to the knowledge at his command; that he will make intelligent decisions and choices in respect to the goods he buys, the business he conducts, the employments he selects, the loans he makes. When we say that men act intelligently, we mean that they take into consideration the known facts relating to the market, that they compare goods and prices, and that they follow the course of action that leads to the maximum gain as they see it. We do not mean, as economists have sometimes been supposed to believe, that men are mechanical calculating machines constantly engaged in weighing pleasures and pains. We assume that since men are in business they will try to make the largest possible ultimate gain for a given outlay of money and effort. We know, of course, that none of these assumptions are true of all people at all times. This economics does not assert. It does insist, however, that the assumptions are true of the great majority of men in their economic activities.

UTILITY

All things that a consumer consciously and voluntarily chooses to buy must have for him some personal significance; that is, at the time of choosing and buying he believes that they will satisfy some personal want. These things are the objects of desire, and they stimulate or produce in the individual the tendency to acquire them. Stated briefly, utility is the capacity of a good to satisfy human desires.

Much needless confusion has arisen in economics over the implications of the term "utility"; and if the student is to avoid this confusion, it is necessary that he should be made aware of its causes at the outset. In the first place, the word "utility" carries no moral significance. If at the moment of purchase a good gratifies a desire, it has utility, no matter whether it confers a lasting benefit upon its buyer or inflicts a permanent injury. Thus the narcotic has utility for one who has become habituated to its use, just as it has for the sick person for the relief of whose suffering it is prescribed by the physician.

In the second place, the want that is satisfied by the good possessing utility is not necessarily a sensuous one. We desire many things that

do not give us pleasure in the narrow meaning of the term. We want books, not only because they amuse or delight, but because we believe that they will instruct, because the habit of reading has been developed in the schoolroom, or even because a friend has suggested to us that we ought to read them. We frequently wear a certain kind of clothing because it is the fashion or the custom or because we wish to avoid the criticism of others who expect us to wear it. Nevertheless these goods are the means by which we satisfy a want or accomplish certain desired ends.

In the third place, we must keep in mind the fact that utility pertains to individual reaction. It is the power of a good to satisfy the wants of an individual, and not the abstract and general significance of a commodity for the comfort or well-being of the community at large. This statement does not deny that individuals are controlled by the action and opinions of larger groups. They buy goods because their friends buy them and because they are subjected to the same advertising, the same climatic environment, the same general social influences as their friends. With all the similarity of reaction to these stimuli there are, however, individual variations that are of importance to economics. Hence we must refrain from using the term "social utility" to mean the want-satisfying power of goods for individuals taken collectively as members of society.

The utility of a good is indicated and measured by the individual reaction to that good. The particular qualities that cause the good to produce in the mind of the consumer the desire to acquire it are controlled partly by his biological make-up, partly by environmental conditions, and partly by acquired modes of reaction. Bread has utility for a large portion of the race because it satisfies a biologic need. The utility of fur coats in winter and of light clothing in summer depends upon changing conditions of the weather. The utilities of many other goods, including things that serve no obvious physiological need, are controlled by the social environment rather than physiological need. Our notions of what is beautiful in the decoration of our houses, of what is desirable as recreation, or of what is in good style in our clothing are controlled very largely by the opinions, customs, and fads of the community in which we live. Utility is not innate in the good itself, nor is it determined by a changeless attitude of the consumer toward his natural environment. Men are not furnished at birth with ready-made utility curves. Education, the appeal of advertising, the dictates of fashion, all operate to modify the utility that goods have for consumers.

Finally, the term "utility" can be strictly applied only to goods intended for personal consumption by the buyer or by someone for whom he is acting. Thus cloth has utility because it ministers to a human want, but the machine on which it was woven has none. Raw materials, the larger proportion of land, and other producers' goods have no specific utility. It is sometimes said that such goods have derived utility because they aid in producing finished goods that have the power to satisfy wants. But this terminology may lead to confusion. It can hardly be asserted that a machine used in fashioning one of the parts of an automobile has utility to the users of the motor cars it aids in turning out. The users are generally unaware that such a machine exists; and even if they do know about it, they do not desire it. To them it is a matter of indifference what machines or what materials are used to make a car, as long as it satisfies their desires. It is also inaccurate to say that the machine has utility for the owner of the manufacturing plant. He may "demand" it because it performs certain work which in turn enables him to make a profit on his investment. But the basis of his demand for the machine is a cost-profit calculation and not a calculation involving his personal satisfactions. Clearness demands that we should not apply the same term, "utility," to two different phenomena. It is better simply to say that the machine is of economic significance to the manufacturer. He wants it because it aids in the production of goods that he expects to sell at a profit.

UTILITY AND SCARCITY

It is a matter of common observation that if goods can be had for the mere trouble of taking them they have no value and no price. But this condition cannot exist unless these articles are so abundant that all may consume them at will. Such objects are called free goods. On the other hand, all goods that do not exist in large enough quantities to satisfy all who desire them are called economic goods. Both types have utility, but only the latter has value. The relation of scarcity to value seems at first glance a simple matter. The occurrences of everyday life lead us to conclude that unless a good has both attributes, scarcity and utility, it cannot have value. For many years economists were content to stop with this conclusion. But this statement did not explain why some goods are more valuable per unit than are other goods, why the consumer is willing to pay a high price for a yard of silk and only a low price for a yard of cotton, or why high prices for both goods prevail at times and

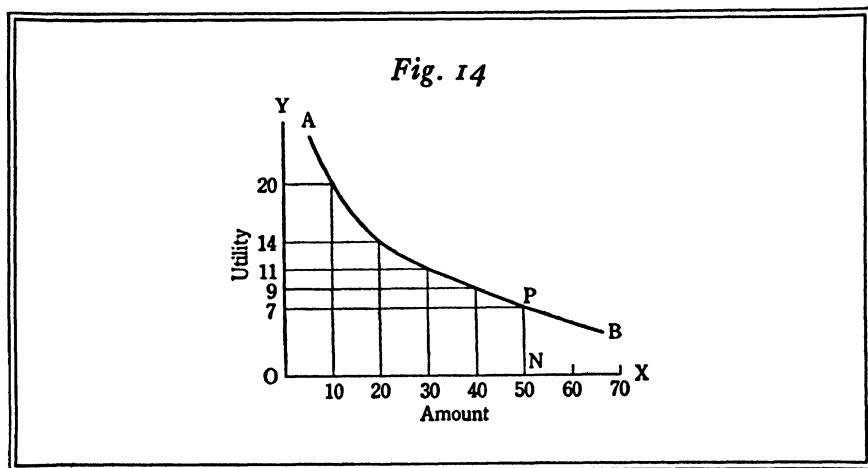
low prices at other times. It is true also that a brief experience in selling any commodity will soon convince one that, as a rule, increasing amounts can be marketed only if the price is lowered. But the market does not offer any explanation of this inverse relation between price per unit and quantity marketed.

THE PRINCIPLE OF DIMINISHING UTILITY

The inverse relationship between price and quantity brought to market depends in part upon the principle of diminishing utility. Since utility is an individual matter, we must begin our explanation with the reactions of a single individual who acquires varying amounts of a consumers' good. Let us assume that the good is a necessity of life, such as bread, for which there are substitutes, but none entirely satisfactory to those who have become accustomed to its use. If the quantity annually produced were much more limited, many persons would be compelled to curtail their use of it. And if the quantity available for consumption were gradually diminished, the loss would be felt more and more by each individual. With present abundance no one attaches much importance either to the acquisition or the loss of such a small quantity as an ounce. But if everyone were reduced to a siege ration of one or two loaves a week, the loss of any given ounce would mean a sensible deprivation. Let us suppose that each person has only four ounces a day for consumption. If one ounce more is added to his allowance, he will feel a certain added gratification, but will find that his desire for bread is more nearly satisfied. If a second ounce is given him, his satisfaction will be still more nearly complete, and the second ounce will not seem so important to him as the first that was added to the existence allowance of four ounces. The utility of a unit of bread is thus seen to depend upon how much is available for consumption during a given period of time. The want that any given unit satisfies is more or less intense, depending on how many other units have already been consumed or are available for consumption. And the relation between utility per unit and quantity is seen to be an inverse one.

Without asserting at this point that the principle is generally applicable to all consumers' goods we may state the law of diminishing utility in respect to such goods as follows: *As the quantity of a good acquired for consumption increases, the utility of each unit acquired is less than that of the preceding unit.* Strictly speaking, this is not a law of *diminishing utility* but of *the inverse relation of utility per unit and*

amount available for consumption. This relationship may be illustrated by the curve *AB* in Fig. 14. Along the axis *OX* are measured the amounts available for consumption; and along *OY*, the utility per unit. If only ten units are available, the utility will be high; but if more units are added, the utility of each will become less and less. In the figure, definite arithmetical values are given to the ordinates extending from *OX* to the curve *AB*. This involves inaccuracy, because there is no method of measuring utility except in terms of something that the in-



dividual will sacrifice to obtain a given commodity. The absolute amount of utility perhaps might be measured accurately if we were able to gauge the varying mental reactions of the consumer by some purely objective standard, as we can his purely muscular reactions to an electric current of varying intensity. Yet the measurement of his response to the good in question in terms of money is certainly not grossly inaccurate. The amounts of money a consumer spends for any one good (not an entire class of goods, such as food) is, as a rule, so small a part of his entire income that the expenditure does not make each dollar more significant than it was before. As Alfred Marshall put it, we may assume that the utility of money remains practically constant. If now the consumer is willing to sacrifice less and less money for units of bread as he is supplied with more and more of it, the inference that this is true because to him the utility of bread is declining cannot be much in error.

The proof of the law of the inverse relation of utility and amounts available for consumption rests primarily, then, upon the observed fact

that as the individual extends his consumption of a commodity his willingness to pay money for each unit declines. Moreover, statistical studies of the demand for goods used almost entirely for human consumption, such as potatoes, show a pronounced decline in the amounts of money people are willing to pay for a unit when the total amount to be sold is increased.

It is sometimes asserted that the principle of diminishing utility applies only to dinner-table examples. But this criticism is without foundation. The use of all consumers' goods, whether articles of food, or clothing, or housing, or amusement, exhibits this same phenomenon of variation in utility per unit as the quantity varies. There are, however, certain limiting conditions attaching to the statement of the law.

(1) A definite period of time must be set within which the consumption of the good takes place, or within which its use is contemplated. If the individual consumes six ounces of bread today, an additional six ounces acquired for use tomorrow will have the same utility. This the law does not deny. It only asserts that if he adds another ounce to his daily ration this additional quantity will have less utility than the preceding amounts. (2) The individual must retain the same attitude toward the good in question. If his tastes change, additional units of the good may come to have either a larger or a smaller utility. For example, a person may care very little for concerts, but, owing to some accidental circumstance, he listens to several under conditions that enable him to enjoy them more than he did before. Because of these changed conditions his enjoyment increases, and, instead of buying a ticket occasionally to occupy an otherwise dull evening, he becomes a regular patron of music. Other examples, particularly those following from the use of any habit-forming stimulant, will occur to the reader. (But in all these cases it will be found that the attitude, the psychological background, or even the physical condition of the consumer has changed.) (3) There must be no change in the environment which will affect his desire to consume. With the change of seasons from autumn to winter, the outdoor worker finds that the utility to him of a given additional quantity of food increases.

MARGINAL UTILITY

As the amount of a good acquired by the consumer increases, the utility of each unit declines. If the stock (Fig. 14) is limited to fifty units, the utility of the last unit will be represented by the ordinate NP

Since the fiftieth, or last, unit has a given utility measured by this ordinate, that degree of utility is called the final degree of utility, or the marginal utility, of the entire stock of fifty units. Since the units are, by assumption, interchangeable, no unit of the stock can have either a greater or less utility than this last, or marginal, unit. To prove this statement let us assume that one of the units other than the last one has been lost or destroyed. How great will be the deprivation suffered by the consumer? If the units are alike and interchangeable, he can put the last unit in the place of the one lost. Hence the loss is measured by the utility of the last, or marginal, unit. The effective importance of the entire stock is measured from the utility or importance of the last unit; what is called the *effective utility* of the entire stock is found by multiplying its marginal utility by the number of units consumed or acquired. The effective utility of any stock is distinguished from its *total utility*. If a monopolist could dole out to the buyer the stock of fifty units one by one and at the same time keep him in ignorance of the existence of any further supply, then each unit would be appraised by the buyer at the utility measured by the corresponding ordinate extending from OX to AB . The first few units would have an extraordinarily high utility, the tenth a lower utility, the thirtieth a still lower utility, and so on. The sum of these utilities would be measured by the area between the axis OY and the ordinate NP and between the curve AB and the axis OX . Total utility is evidently a less definite and measurable quantity than effective utility. No one knows how great is the utility to him of a very small quantity of many of the commoner necessities of life, and to construct a curve showing the total utility of such articles is therefore impossible. With less necessary articles of consumption the case is somewhat different. Most of us are quite aware that we would not buy ice cream at five dollars a pint, because the utility of even one pint is less than the utility of many other things five dollars would purchase.

Thus far we have considered the application of the principle of diminishing utility to goods of which the consumer buys a large number of interchangeable units. But many things are not of this sort. We do not buy unit after unit of pianos, or houses, or automobiles. Do the phenomena of diminishing utility and marginal utility apply to such goods? Commodities such as bread satisfy very simple desires. Other goods, such as automobiles, cater to complex desires. The latter good may be desired because it provides comfortable transportation, because it furnishes the owner with a kind of recreation, because it ministers

to the desire for distinction, and for many additional reasons. Now the utility of a good depends, of course, upon the services it renders, upon the desirable events that, in the opinion of the user, flow from that good. This is true of automobiles as well as of bread, but in the case of bread the service is brief, the commodity quickly consumed, and the want simple. In the case of the automobile the services are numerous and continued over a considerable period of time. Hence the buyer of the latter good usually considers the desirability of each of the several services and also the repetition of those services. The principle of diminishing utility applies to the repetition of each of the services of the automobile. For example, a consumer might well ascribe a high utility to a small amount of one of the various kinds of service rendered by an automobile, but more than a small amount might have for him a very low utility. If he greatly desired a small amount of automobile service,—say, rapid transportation,—he would probably hire a taxicab rather than buy a car outright; but the phenomenon of diminishing utility would still apply to his consumption of this service. Because the service derived from bread is quickly rendered, the identification of the service with the physical object is close and we do not observe that the two are different. But in the case of the motor car the difference is easily seen. In both, however, the utility is derived from the use of the good, and a given amount of service from either would have, for a particular consumer, a given marginal utility.

UTILITY AND THE INDIVIDUAL-DEMAND SCHEDULE

It is the purpose of this section to explain how a rational consumer may arrive at the demand prices he sets upon different amounts of the various goods he buys. A great majority of consumers do not receive incomes sufficient to gratify all desires. Many objects in a consumer's environment attract his attention and produce in him a reaction that disposes him to use or possess them, but he has not the wherewithal to acquire them all. Since the consumer cannot gratify his every whim, *he must make choices among goods*, acquiring some of them and forgoing others. For purposes of illustration let it be assumed (1) that there are only three goods among which an individual at a given time can make choices, and (2) that he has not the means to acquire all he wants of all three. On what basis will he make his choices? Which of the three will he acquire? Or if he acquires some of each, what will govern the amount of each that he acquires? The answer to these

questions is that if the consumer is rational and carefully considers the utility to be obtained, he will make such a selection among the goods as will give him the greatest satisfaction. He will be attracted first to the good that has for him the greatest utility. If his income is very limited, he may spend it all in acquiring this good and nothing else.

In order to explain the choice of different amounts of several goods we may suppose that the consumer has the power to buy 100 units of bread, or 25 units of meat, or 20 units of pastry. But if he acquires much bread, he will find that its utility declines as the amount is increased. The hundredth unit may have very little significance, and consequently the total stock will have a small effective utility. Under the circumstances he will turn to meat and even to pastry before he acquires 100 units of bread. As he increases his stock of each, the utility of an additional unit of each will decline; and when he has entirely expended his means of acquisition, each good will have for him a given marginal utility. If he is rational, and if he carefully considers the marginal utility of each good, it will be found that when he has expended all his income he will have so distributed his consumption that, *relative to the prices he has to pay for the various goods*, his marginal utility for all will be equal. The marginal utilities will not be *absolutely* equal. If the prices of bread, meat, and pastry are 10, 40, and 50 cents a unit respectively, then the marginal utilities of the three goods will stand in the relation of 1, 4, and 5. The consumer would not continue to buy meat at the point where its utility, relative to bread, was as 2 to 1, because he would then be spending money upon meat that would procure for him a greater satisfaction if expended for bread. He would not continue to consume pastry to the point where its marginal utility was only four times that of bread, because at that point he would be compelled to pay for a unit of pastry 50 cents, which would buy an amount of bread having more utility per unit. The prices the consumer tends to pay for goods are thus seen to be proportional to the marginal utilities these goods have for him.

If we bring the illustration nearer to real life by introducing many commodities, the explanation will be in no wise altered. Instead of comparing the marginal utilities of three goods, the consumer would seem to be confronted with the problem of comparing the utilities of a large number of goods. But in deciding whether to buy a given quantity of a certain commodity he does not need to make the multitude of comparisons suggested in the preceding sentence, provided (1) money is in common use and (2) he has established a scale of expen-

diture in which he can place with more or less accuracy the utility of any good with which he is familiar. He knows that if he buys too lavishly of one good, he will have to go without a number of other things that would have for him a greater power to satisfy wants. He tries, therefore, to distribute his income so as to obtain the maximum return in utility. In a figurative sense, money comes to have a definite utility per unit, because it comes to stand for a composite of utilities derivable from a large number of goods. The consumer therefore tends to continue his purchases of a particular good until the marginal utility he obtains from it is equal to the utility represented by the amount of other goods he could purchase with the money the last unit of the particular good costs.

In the case of durable objects, the consumer compares the succession of uses, pleasures, or benefits he will receive from them during a period of time with other satisfactions their cost in money would buy. Here again he will tend to lay out his income in such a way that the money spent for such goods will yield him about the same marginal utility as could have been purchased if the money had been used to buy other commodities. The preceding statement is subject, however, to the qualification that the utilities to be derived from the use of a permanent good must ordinarily be greater than the utilities of an impermanent good having the same price. This is true because the satisfactions derived from the durable good are deferred in part to the future, and it is a common rule of life that future satisfactions are undervalued as compared with present satisfactions.

In the preceding discussion the reservation "if he [the consumer] carefully considers the marginal utility of each good" has been employed. The meaning and significance of this reservation must now be explained. If the buyer should purchase at random, paying any price asked, or if he should never stop to consider what any good cost him in terms of other goods given up, or if he should act blindly and on mere impulse, then for economics his conduct would be irrational. Under these conditions his choices of goods would have nothing to do with utility. In place of making his choices on the basis of comparative utilities, he would make them on no basis whatever, and utility would be irrelevant. But under the conditions of buying and selling as we find them the consumer does make choices, and he does weigh the desirability of one good, or a given number of units of one good, against varying quantities of other goods. The proof of this statement is found everywhere in the market. The art of consumption is far from being

exact, but the practice of it is not as chaotic as is sometimes represented. Automobiles that do not ride comfortably cannot be sold; stoves that do not heat well are discarded; clothing that does not wear well or that proves unsatisfactory in other respects remains on the shelves of the dealers.

In some instances the consumer may be misled. The expert can show that he has been deceived and that he really chooses the inferior good. But this is beside the point. Regardless of the wisdom of the consumer's selection, he has made a rational choice. It has been objected also that the consumer cannot compare the utility of literally thousands of articles offered for sale in the shops of a large city. This is partly true. It is especially true in the case of goods that the consumer buys only infrequently and concerning which he has therefore only a limited experience, or in the case of goods that are new to him. In these instances the shrewd buyer will usually do one of two things: either he will patronize dealers who have established reputations for honesty and fairness, or he will try to obtain the opinion of someone who has had more experience—the expert, perhaps. Thus one buys a new accessory for his automobile on the opinion of others who have used it, on the advice of a mechanic, or on the recommendation of a reputable dealer. In any case, it is clear that in this instance the user of the article is partly substituting the choice of another for his own judgment; but only partly so, for no matter how highly recommended the good may be, the prospective buyer may still conclude that he cannot afford to buy it. That is, he decides that, on the representation of other people, the utility of the good is not great enough to warrant the expenditure of the money necessary to obtain it. The major part of one's income is not expended for goods with which one has had no experience; in most cases the buyer has used the identical good or has used one sufficiently like it to provide a basis for comparison.

Goods commonly consumed, however, are so numerous that the buyer cannot stop every time he makes a purchase to compare its marginal utility with the marginal utility of money or with the utility of the other things its price represents. Fortunately he does not need to do this. Any experienced consumer soon establishes a scale of expenditure in which he has placed roughly the relative importance of quantities of various sorts of goods. The places of different goods in this scale depend on a great number of considerations: income, education, the acts of associates, climate, advertisements, and many other conditions. The net result of this ranking is that the consumer comes to know how much a

dollar is worth to him in terms of given quantities of a great variety of goods. He knows without elaborate calculation that if he buys a pair of shoes for double the price he is accustomed to pay, or if he buys two pairs of shoes instead of one pair, he will be compelled to go without certain other goods to sacrifice which will cause him annoyance or disturb unpleasantly the routine of his life. If the purchase is an important one, he will stop to consider these sacrifices. He will weigh as well as he can the various returns in satisfaction he might obtain from each good. But he will continue to make numerous small purchases at approximately the usual prices and in the usual amounts without giving either much thought.

In the case of goods that cannot be bought in small units, or whose services cannot be obtained in small units, the consumer may find it impossible to arrange his budget of expenditures so that the marginal utilities of all the goods he purchases will be proportional to the prices he pays. For example, a person wants a radio for the purpose of listening to programs of a certain sort (say, reports of football games), but he may not attach a high utility to any additional use that he could make of it. He wants the equipment for use on only six or seven afternoons during the year. For this service he is willing to pay \$10 a year. But he finds that the annual cost of such an outfit as he would require is \$20. The additional services he could get from it are not worth \$10 annually. What should he do? If he buys the radio, he will have to forgo some service or commodity that has a cost of \$10 and a higher utility than the additional service of the radio. If he does not buy the radio, he will probably spend \$10 on some service or commodity that will not give him as much satisfaction as he could obtain from listening to the reports of football games. He may be able to strike a balance between the excess of satisfaction derived from the reports of games and the loss of alternative enjoyments resulting from paying \$10 for the additional radio services and giving up other more important satisfactions. But if the broadcasts of games are worth only \$10 to him, he will not be able to do this. Then he will have to make up his mind to losing potential satisfaction either way he decides. He will be unable to derive the same satisfaction from marginal expenditures on all the commodities and services he buys.

The individual-demand schedule represents the prices that an individual stands ready to pay for different amounts of a given good. From the preceding explanation it will be seen that these prices will be paid by the individual only if (1) his money income remains the

same, (2) his demand schedules for other goods remain the same, and (3) his desire for the given good remains the same. If his income shrinks, he will need to revise his scale of expenditures for all goods. Hence he will not be willing to take as much of any good at a given price as he would have taken before his income declined. Likewise, if the marginal utility of one good is increased by a change in tastes, the consumer will desire more of it and will expend more money for it. This will necessitate a reduction of expenditures for other commodities, and therefore the consumer's demand schedules for all these other commodities will be altered to some extent.

MARGINAL UTILITY AND MARKET PRICE

As we have seen, the marginal demand prices of an individual for different goods are, in the main, proportional to the marginal utilities of the goods to him. But the question arises whether the market prices of different goods are proportional to the marginal utilities of these goods to all the buyers. The answer is that it is possible, but unlikely, that this proportion would prevail. In the first place, because of differences in income, the prices that different consumers stand ready to pay for the same quantities of the same good are unlike, and the amounts that they buy of the same commodities at the same prices are unlike. The marginal utility of money is higher for the people of low incomes than it is for those who receive higher incomes. Hence the purchase of a ticket to the theater at three dollars by two persons who have different incomes may be a marginal purchase for both and yet not represent the same satisfaction to both. For the one who has a small income three dollars may represent a commodity or a quantity of commodities that have a high marginal utility, but for the other it may represent goods that have very little marginal utility.

In the second place, we know that the rich buy and consume many commodities and services, that the laborer's family never use at all, and that, to a less extent, the latter buy and use commodities that the rich do not consume. Between these two groups of commodities, a unit of each of which may sell for the same price, there is no necessity for equality of marginal utility any more than there is between two commodities sold and consumed exclusively in different parts of the world. The users of the commodities are divided into noncompeting groups. In somewhat the same way, extreme differences in tastes and habits of consumption may separate two groups who live in the same

city and have the same incomes. There are commodities that are highly prized by some—have high marginal utility—and not consumed at all by others because of these differences. And in this case it does not hold that the market prices of goods are proportional to their marginal utilities. The marginal buyers of different goods have not the same scales of satisfaction in respect to different goods.

SOME PRACTICAL APPLICATIONS OF THE MARGINAL-UTILITY THEORY

The utility theory, which includes the principle of diminishing utility, is the only reasonable explanation of the fact that many of the necessities of life sell at low prices and command only a small portion of the income of even the wage-earner. The world could ill afford to do without bread, and yet many wage-earners' families expend much more income for comforts and decencies than for bread. Of course the answer to this apparent riddle is that the amount of wheat that can be obtained from the grain lands of the world is so great that its marginal utility is low, whereas the sources of less necessary goods are more limited by nature and by the disutility of labor and saving, and thus their marginal utilities are high. The utility theory also explains satisfactorily the decline in price when the amount of any good offered for sale in the consumers' market is increased.

One of the most consistent tendencies of modern economic life is the constant expansion of wants. As invention and scientific discovery increase man's power to produce, we find that, instead of confining production to old commodities, he constantly invents new ones. Both production and consumption become more and more complex. Part of this complexity is due to the fact, stressed by Alfred Marshall, that new activities create new wants. It is due also to the principle of diminishing utility. As the old commodities are multiplied, they come to have less and less want-satisfying power; and as their utility declines, the producers, who are ever on the alert to sell goods, introduce new ways of satisfying old desires and new modes of consumption.

The effectiveness of advertising in expanding sales and in keeping up prices rests on its power to bring goods to the attention of potential consumers who otherwise would either not know that the goods existed or not have their attention constantly directed to them. In the latter case, the advertisement works to produce a constant comparison between the want-satisfying power of the good to be sold, often at the

advertiser's appraisal, and the want-satisfying power of other things its price would buy. Here again the utility theory furnishes a hypothesis upon which we can explain the observed facts.

The utility theory is helpful also in such practical matters as the distribution of the burden of taxation among different income groups. The rich are justly taxed more heavily than the poor for the obvious reason that the poor cannot pay. The rich are taxed more heavily also in proportion to their incomes than the lower middle class because it is reasonable to suppose that a deduction of 20 per cent from an income of \$100,000 deprives the recipient of a smaller sum total of utilities, or causes no more sacrifice than a deduction of 10 per cent from an income of \$5000. As will be shown later (Chapter XXXII) this principle cannot be extended too far. It scarcely applies when we differentiate between incomes of \$100,000 and \$500,000. The utility theory also applies roughly in the various measures for the relief of certain classes of society. If it is assumed that all incomes are partly the product of the environment in which all the people of a community live, then to take something from the rich (by taxation or voluntary contribution) to relieve the aged, the indigent sick, and those who are involuntarily unemployed is to increase the general well-being of the community. The funds thus employed add more to the utilities enjoyed by one group than they subtract from those enjoyed by the other group.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It is sometimes said that the utility explanation of demand price is faulty because custom, whim, and the persuasive arts of salesmen determine the prices consumers are willing to pay for goods. To this we may reply: Economics does not inquire how the consumer got his desires, but only shows how he makes his choices, desires being given. Do you think the reply is adequate?

2. "Utility has nothing to do with demand price. If it did, people would cease patronizing the 'movies' and go to symphony concerts, refuse to buy gaudy, un-serviceable clothes, and insist on quality rather than appearances in their purchases." Show that in this quotation an erroneous definition of utility is implied.

3. The utility theory is frequently criticized because, it is said, it is based upon untenable assumptions. Which of the assumptions enumerated in the latter part of the first section of this chapter do you believe are least tenable?

4. One writer on the theory of value has said that the utility of any good is regulated by the utilities of other goods. Show that this is partly correct, but that it is not correct if applied to our desires for certain types of goods. To which types does it apply? not apply?

5. The operation of the principle of marginal utility is sometimes interfered with because the objects the consumer desires are not divided into small units. Enumerate commodities that you think might be sold in larger amounts if they were salable in smaller units.

6. In this chapter, it is said that the reason why people will buy increasing quantities of goods only at declining prices is the principle of diminishing utility. Assume that this explanation is incorrect. What better explanation can you suggest?

7. "The principle of diminishing utility is a mere assumption. It is inferred from the fact that prices decline as the quantity of goods offered is increased." Is this statement correct? What other evidence of the existence of the principle of diminishing utility can you give?

8. A thousand families having widely different incomes all buy milk at the same price. Is it possible that all these families are marginal buyers for the last units purchased? If so, have the quantities taken by each family the same marginal utility?

9. The elasticity of the collective demands for different goods is partly explained by the principle of diminishing utility. Why? What other economic facts affect elasticity?

CHAPTER X · Cost of Production and Normal Price



HOW SUPPLY IS DETERMINED

The market price of any good is determined by demand and supply. Demand, as we have seen in the preceding chapter, is the result of the great variety of conditions that determine the marginal utility of a good to consumers and the purchasing power that the consumers of the good can command. Supply is determined by the decisions of the sellers, who set reservation prices upon products and offer them for sale at the prices they believe will yield them the maximum gain. The conditions affecting the reservation prices that sellers place on goods have been stated. Among these conditions, as was stated, is the cost of production of the goods.

It is evident that if the different amounts of goods were arbitrarily fixed by the government or some other agency, the price paid for each commodity and the relative prices of different commodities, or the exchange values of all commodities among themselves, would be determined by demand alone. On the other hand, if the demand schedules of all goods were fixed and invariable, the amount of each commodity offered for sale would determine its price, and the relative prices of different goods would be determined by the relative amounts thrown upon the market. Now the relative amounts of different goods put up for sale are not determined arbitrarily any more than the demand for them. It already has been explained how the forces behind demand operate, and, in part, how the supplies of goods that either are completed and ready for sale to consumers, or at least have been completely manufactured, are regulated. But when a good has reached the retailer it is nearly in a class with land. Certain quantities are already in existence, and, unless the prices that consumers offer will not compensate for the cost of retailing, the goods will be sold for whatever they will bring. But the explanation of how the supplies of reproducible goods come to be what they are, which is an extremely difficult problem, has not been touched upon. This explanation involves the theory of cost in relation to price.

COST OF PRODUCTION AND PRICE

The price that is set on finished goods when they are thrown upon the market to sell for whatever they will bring is a temporary equilibrium price. The forces of demand and supply operative in this situation are relatively simple. On the demand side are the forces explained in Chapter IX. On the supply side the only "force" is the quantity of goods in existence. But, as everyone knows, the supplies of goods that flow onto the market do not remain constant. They are increased when producers find it to their advantage to augment them and decreased when these producers believe they can make greater profits by reducing them. The environmental conditions that influence producers to increase the supply of now one good and now another and to cause supplies to decrease are multitudinous. This text will not attempt to explain all of them. In fact, it is probable that not all are known or, if known, not all are understood. Moreover, these conditions change with great frequency in any modern industry. A description that would account accurately for the supply of wheat during 1935 would probably not entirely account for the supply of 1937. To attempt to set up a model description of the conditions affecting the supplies of all goods would be futile. This we shall not attempt to do. Instead we shall first explain how the supplies of goods would be regulated under what were formerly called "static conditions" and what are now more frequently called "normal" or "equilibrium conditions." This explanation will be the subject of this and the following two chapters. Chapters XIII and XIV will give an introductory explanation of the effects of certain non-static or disequilibrium conditions.

To make an explanation of price under equilibrium conditions it is necessary to lay down a number of assumptions. These assumptions are more or less unrealistic and some of them will be withdrawn in later chapters. In the first place, we shall assume that competition between all sellers as well as between all buyers is free and unhampered. Every producer can quit the line of production in which he is engaged and move to another without loss of capital or any other asset. He also can sell in any market without let or hindrance so long as he obeys the laws regulating contracts, the possession of property, and the personal rights of others, and so long as he keeps the peace. No force can be used to prevent his seeking his maximum profit either by individuals or by the state.

In the second place, we shall assume that in every market there are

either actually or potentially so large a number of sellers and buyers that the withdrawal of any one of them will not make any appreciable difference in the supply or the demand. We see this condition of supply in the United States market for wheat at the present time. There are literally millions of wheat-growers, and no one of them can possibly change the quantity offered for sale in any year by withdrawing his entire crop from sale. Neither could he affect it by increasing his crop to any amount that is within his power. Each producer is, as it were, a drop of water in an ocean. Whether he is there or not makes no measurable difference in the sea level.

In the third place, we assume that the quantities of the primary factors of production are not changeable. The quantities of land in use and of capital in existence, the numbers of laborers at work, and the numbers of entrepreneurs engaged in performing the managerial function are assumed to remain the same.

In the fourth place, the demand of consumers must remain fixed, and, fifth, the techniques of production also must remain as they are without any progress or retrogression. Finally, we assume that labor and liquid capital are free to move from place to place and industry to industry as easily as the entrepreneur.

Under these assumptions it is easy to see that after a time all the factors of production would come to earn the same return in all their employments. That is, labor of a given grade of natural ability and training would earn the same wages in all employments. Land of a given fertility and location would earn the same rent no matter what it was used to produce, and capital would earn the same interest or profit in all enterprises. After these equalities had been attained we should have an equilibrium. That is, no factor of production would be changed to any other use because no gain or profit could be made by changing it. The remainder of this chapter deals with the determination of price under these conditions. This price is usually called *normal price*.

If we should ask a manufacturer of (say) cotton cloth what he considers a normal price for the commodity he produces, he would probably reply that a normal price is one that covers his payments to other people for raw materials, labor, borrowed money, land rent, and so on, plus a fair return on his own investment and for his own work devoted to the management of the business. Of course he would say also, if he were quite frank with us, that he would like to sell at a price considerably above these costs of production in order to earn a very high

return on his investment and for his own work, and that he would take advantage of any temporary shortage of supplies to charge the highest price he could get. But he would add that when such fortunate opportunities did come to him they were likely to be offset by the losses of bad years, so that, in the long run, he could not expect to receive a price for his product in excess of cost.¹

If we should next ask what determines the rates at which our hypothetical manufacturer must buy the factors of production from other people, or what determines his out-of-pocket expenses, we should be answered that, as far as the single producer is concerned, they are determined by the competition of other producers who wish to employ the agents in producing cloth. And if we should ask what determines fair return on the invested capital of the manufacturer, we should learn that each investor expects to obtain the competitive return for investments of equal risk and that no one puts his funds into a business unless assured that in all probability he will receive this return. Hence we should find that the costs of the factors of production are determined proximately by the competition of entrepreneurs within the same line of production.

But the land on which cotton is grown could be used also for the production of several other crops, the capital invested in machinery could be used in other businesses, and the labor used in the cotton mills might be employed in some other occupation. Evidently, then, it is not the demand of the cotton mills alone that sets the prices of these agents. The cotton industry must bid against many other industries for land, labor, loanable funds, entrepreneurship—in fact, all the factors of production. And the costs of the factors of production for making cotton cloth are thus set by the competitive bidding of all industries using the same factors of production. Hence we conclude that the total cost of production of cotton cloth is determined by the value which entrepreneurs in all industries attach to the factors for producing cloth. Costs thus determined are called *opportunity costs*. }

The objection may be made at this point that we have not explained how the prices of the factors of production, land and labor, are ultimately determined. We have said only that the prices of the different finished commodities tend to the level at which the factors would earn the same in all industries and have not given any conditions that fix the values of these factors. It must be admitted that no explanation

¹ The student's attention again is called to the fact that "cost" includes a fair return on the entrepreneur's own investment and labor.

of the values of these agents has been given. But for the present discussion this explanation is not required. The ultimate determination of the values of the different groups of agents, the factors of production, is not required for the present discussion. We are here concerned with finding out how the costs of production of the products of a single industry or of a single commodity are determined. Producers in a single industry cannot affect these costs very much. They cannot, as a rule, affect the total demand for the agents of production enough to cause a change in their prices.

Later the explanation of the values of the factors will be considered, and it will be shown how these values are determined.¹

Still another assumption must be made at this point. We know that while the factors of production tend to have the same costs to all entrepreneurs in the same market, the average unit cost of turning out a product is not the same in all plants. Some entrepreneurs are more skillful than others in managing production, and, in consequence, average unit cost varies from plant to plant. This difficulty may be put aside for the present discussion by assuming that entrepreneurs are of equal efficiency, and that total unit cost is the same in all plants.

NORMAL PRICE

Since the unit price of any commodity tends toward its unit cost of production, it is sometimes said that normal price is "cost price." This statement is correct, but it throws very little light on the forces determining the norm about which market prices fluctuate. In the following explanation of normal price, we shall resort again to the concrete illustration of cotton cloth. We may put the problem as follows: At what price will the supply of cloth from all the mills be such that, in the absence of temporary disturbing conditions causing market price to fluctuate about normal price, the demand for cloth will just absorb the entire annual product and no more?

The amount of cloth that can be sold at high prices is less than the amount that can be sold at low prices, because the demand curve for cloth, like all other demand curves, slopes downward. Hence, if the output of the mills were increased gradually, the price presently would fall below the cost of producing cloth. This would not be a profitable condition, and sooner or later some labor and capital would be shifted

¹ Chapters XXIII-XXVII, on wages, rent, interest, and profits.

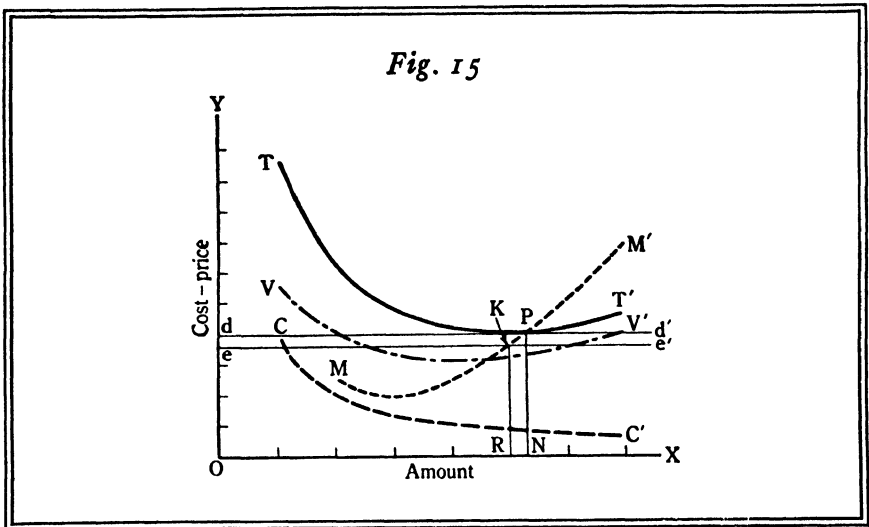
from the cotton mills to other enterprises. On the other hand, if the amount of cloth were gradually reduced, a time would come when the price was above the cost of production, and the mills would make a larger money gain from the use of the agents of production than entrepreneurs in other industries could realize from the employment of these same agents. This would cause a drift of labor and capital into the cotton industry, the output of cloth would be increased, and price would fall.

At some point in the upward and downward movements of the output, however, there would be a price that would neither invite any new labor and capital to enter the industry nor encourage any of these agents already in the industry to leave it. This price would be reached when the agents of production employed in cotton manufacture earned exactly the same wages, rent, interest, and entrepreneur's profit as they did in all other industries. At this price the demand for cloth and the amount produced would be in stable equilibrium. No producer would have any economic reason for either enlarging or contracting the output of cloth. The price accompanying such a stable equilibrium of demand and production is the normal price of any commodity. It is called also the *equilibrium price*. The price of any commodity is normal if the agents of production earn as much in the production of that commodity as they do in industry at large, and no more.

Normal price is not a fixed price. Economic conditions change, and the point of equilibrium of demand and production does not remain fixed. The normal price of a pair of shoes may be five dollars in 1937; five years later, owing to changes in the opportunity costs of the agents of production, it may be either above or below that figure. Neither does the term "normal" mean "ethically right" or "socially desirable." The normal price of a quart of milk may be fifteen cents, and yet it may happen that at this price the children of poor families cannot obtain enough to keep them vigorous and healthy. There can be no criticism of the distributor who sells at this price, if it is a competitive price, although there may be serious social injustice in the fact that the children are denied the requisites for healthful living. It may be entirely desirable for the state to step in and provide a cheaper supply of milk for those unable to buy at the normal price.

THE POSITION OF THE INDIVIDUAL SELLER UNDER PERFECT COMPETITION

It has been explained in the preceding sections of this chapter that cost of production and price will agree if there is complete mobility of the factors of production and if there is perfect competition. It has been stated that the costs that make up competitive price are always opportunity costs. In this section the position of the individual producer in the competitive market will be explained. In Fig. 15, Od or NP represents the price set in the market. Since there must be a multitude of



sellers in any purely competitive market, it is impossible for any one seller to change the price by increasing or decreasing the amount he offers to sell. Hence price will always be found on the line dd' .

Suppose that the price line, dd' , lies above TT' (the total-cost curve) at the point of lowest total unit cost, P . The producer now sells at a price that gives him a profit over and above all cost. But if conditions are perfectly competitive, other producers will find that their cost is the same as the cost of the seller in this illustration. All will make a profit. This means that the factors of production earn more in the production of this good than in the production of other goods. The owners of the factors now employed to produce these other goods will withdraw them and use them to produce this good. This will increase the supply of this good and cause the price line, dd' , to sink to Od or NP .

Again, suppose the price line, dd' , to lie entirely below the curve TT' in Fig. 15, in the position ee' . All sellers find that even when they are producing at maximum efficiency their lowest total unit cost is still above the selling price. Now, since costs are the prices set on the factors in other enterprises, it follows that the factors earn less in producing this good than in producing other goods. Under this condition some of the producers will cease producing this good, those who continue will produce only OR units, supply will decline, and price will rise to NP .

Finally, suppose that dd' lies above or below the point of lowest total unit cost for only one or a few sellers of this good, while it is exactly tangent to TT' at P for the greater part of the sellers. If it lies below, that indicates that the sellers who are so situated are inefficient and must either quit the business or increase their efficiency. If the line dd' lies above the point of lowest total unit cost, that indicates that the few whose positions are so represented have some advantage over other producers that caused total cost to drop below the position of TT' for the majority. If this advantage is the use of new methods of production in which all sellers may participate in time, then the cost curves of all sellers eventually will be lowered until they are identical with those of the few who were first to obtain the advantage, supply will be increased by an influx of additional factors, and dd' will move downward until it is tangent to the point of lowest total unit cost for all. But until all have adopted the improved methods of production there will be no equilibrium price. The price will be subject to dynamic conditions.¹ If the advantage is one in which only a few can participate, then a condition of partial monopoly may exist, and the price may not be a competitive price.²

The conclusion of this discussion of the relations of price and cost may be summarized as follows: (1) cost of production means the lowest total unit cost of the entrepreneurs in terms of prices of the factors of production; (2) the normal, or equilibrium, price of any commodity is the price that corresponds to the lowest total unit cost of the producers of that commodity under conditions of perfect competition between all sellers and all buyers and freedom of the factors to move from one use to another; (3) if, for any reason, price happens to rise above or fall below the lowest total unit cost of the producers, new producers come in and increase the supply or old producers drop out and decrease the supply, and, as a result of these adjustments, price will return to equality with lowest total unit cost.

¹ See Chapter XIII.

² See Chapter XII, where the subject of monopoly price is discussed.

PRIME AND SUPPLEMENTARY COST

In the preceding section we have assumed that if, for any reason, price happened not to coincide with the lowest total unit cost of the producers, supply would be adjusted by the coming in of new firms or the dropping out of old ones. But there is another type of adjustment of supply that is important for the re-establishment of equilibrium price when the factors of production are not instantaneously mobile. This adjustment takes place through the increase or decrease in the physical volume of production by the firms already in the field which will not be forced to cease production if price does not decline too greatly and which will add to the supply of the good if price rises a good deal. As a matter of fact, in the real world the short-time adjustments of supply are more likely to be made in this manner than by the appearance and disappearance of firms. Here we are dealing with slightly dynamic conditions.

It matters little what goods may have cost in terms of past wages, past investment of capital, and past cost of raw materials if a given quantity has been thrown upon the market. As we have seen, this quantity will have to sell for whatever the buyers are willing to pay for it. Hence it is only when past plans for production have forecast cost and demand accurately that goods sell for prices that agree with past cost of production. It often happens that entrepreneurs invest too much capital in the buildings and machinery of a certain industry. Such capital goods cannot be worn out quickly, and often they cannot be used for the production of any sort of commodity except the one for which they were intended. If their earnings fall below the gains that could have been obtained if the investment had been made in other industries, the entrepreneurs have no recourse. They must take whatever the capital will earn.

If, for example, too much specialized machinery has been constructed for the production of cotton cloth, the owners cannot transfer it to the shoe industry or any other industry, no matter what are the profits to be made in these alternative employments. Concrete examples of capital thus unwisely invested are found in every community. In every city there are buildings that cannot be let to tenants at rentals that will pay both the operating expenses and the customary return on the money tied up in the buildings. In certain parts of the United States competing lines of railroads are so numerous relative to the demand for their services that some cannot pay dividends and some cannot even pay interest on their bonds.

Costs arising from interest on fixed, specialized capital and from

depreciation are not, in the short run, opportunity costs. In time, of course, capital wears out, and unless the entrepreneurs believe that new investments in fixed equipment in a field that has been depressed will yield a return comparable to that yielded in other businesses the capital will not be replaced. If it is not, the supply of the products of that industry will be reduced and price will be forced up to the point where returns to capital again will be forthcoming.

In contrast with the outlays that have been made for such agents as fixed specialized capital are payments for common labor, raw materials bought as required for use, and other similar expenses. The expenses that the entrepreneur incurs for raw materials, common labor, and the repairs necessary to make good the effects of wear and tear are called *prime* cost of production. They are the variable cost represented in Fig. 15 by the curve VV' . The expenses brought about by putting up buildings, constructing bridges and tunnels—that is, interest and depreciation on permanent capital—together with the salaries of higher employees, are called supplementary cost, which is represented in Fig. 15 by the curve CC' . In general, supplementary cost results from investment in durable, specialized instruments of production; prime cost results from the purchase of human services and goods that can quickly be turned into finished commodities.

Under equilibrium conditions a good must sell for a price that covers both its prime cost and its supplementary cost. If it does not, entrepreneurs will cease to produce it; for even the most durable capital wears out, and if the investment in it has not proved profitable, such capital will hardly be replaced. This seems a conclusion from which there is no escape. Yet it is a conclusion that rests on the assumption that entrepreneurs can foresee what will happen in the future, which they cannot do. Consequently we find that, because of repeated mistakes on the part of investors and entrepreneurs, an industry may continue for a considerable period of time producing goods at costs (the sum of prime and supplementary) in excess of selling prices.

To show that it is often the most profitable procedure for producers to sell goods at something below the sum of both prime cost and supplementary cost, let us assume that the total unit cost of the good is \$5, its prime cost \$3, and its supplementary cost \$2. If the price falls below \$3, the day-to-day expenses cannot be covered, and the producers will cease hiring labor and buying materials that give rise to a good that is worth less than its prime cost. But it is otherwise if the good sells for anything over \$3. All prime cost is now covered, and there remains

over and above such cost something with which to meet interest charges and depreciation on the fixed, specialized capital. If the entrepreneur ceases producing, most of these capital expenses will go on just the same as if his plant were running at full capacity; so he will continue to produce as long as the good sells for anything above \$3. But when the plant wears out, he will not rebuild it unless he anticipates that in the future both prime expenses and supplementary expenses can be met.

Here, however, the element of risk and the necessity for prediction and forecasting again come into play. The entrepreneur who suffers disastrous losses because he cannot cover all his supplementary cost may not be willing to risk further investments. But business conditions are not often stable, and a year or two of prosperity immediately preceding the time when the decision to invest or not invest must be made may so overshadow more remote losses that both the entrepreneur and the bond-buyer may decide again to invest in the business which in the past has not proved successful. If they decide not to invest, others may take a more optimistic view and thus keep the industry overbuilt for a long period of time. Evidently the railroads in the territory west of the Mississippi River were overbuilt for years, especially during the two decades preceding 1900. Yet this did not deter two of them from constructing, at great expense, two additional lines to the Pacific coast when a few years of good profits, following the revival of business in 1898, had put the older lines on a more solid financial basis.

When producers are confronted with a decrease in demand and therefore are unable to sell the maximum output of their plants, it not infrequently happens that they attempt for a time to maintain prices by reducing output either by running only part time or allowing a part of the plants to stand idle. Usually, when this happens, it would be impossible to sell the capacity output for enough to cover prime cost. But output is likely to be reduced before prices have reached so low a level as prime cost. A common reason for this procedure is that to reduce prices would spoil the market and prevent a revival of demand. The prices at which the goods are sold are supposed to have become customary; and if they were reduced, consumers might refuse to buy when at a later period they were put back again to the old figure.

The amount by which each producer should reduce output when price falls is shown in Fig. 15. The dotted line ee' shows that price has fallen below NP , the lowest total unit cost. In the preceding section it was said that if price remained permanently at this level some of the producers would be forced sooner or later to drop out of production entirely.

This statement is correct, but it is also true that before these producers withdrew their supplies other producers would reduce their offerings. The question is, By how much? The price line ee' intersects the marginal-cost curve, MM' , at K . That is, all the output of the producer whose costs are illustrated in the figure up to the amount OR can still be sold at a price that exceeds average variable (prime) cost, but the last unit at R is sold at a price, RK , that exactly equals the additional cost of producing it, as is shown by the fact that marginal cost (on MM') is also RK . The producer therefore should produce exactly OR units. If he produces less, he loses more than he should, because added units could still be produced for less than they sell for. If he produces more, he loses more than he should, because all units beyond OR add to cost more than they sell for.

We now conclude that when price falls below lowest total unit cost the supply will be reduced for two reasons: (1) some firms will cease to produce; and (2) all producers will reduce output. By similar reasoning, it can easily be seen that when price rises above lowest total unit cost not only will new producers come in after a time, but, in the interim, old producers will expand output. They will expand output to the point indicated by the intersection of the new price line (above and parallel to dd') with the marginal-cost curve MM' .

Whether the conditions that compel producers to sell their products at prices below the sum of prime cost and supplementary cost are permanent or only temporary depends on a number of considerations. In the depression stage of the business cycle many, and sometimes all, producers in a trade are forced either to sell below total cost or to close down. The situation is only temporary for a majority. With the return of prosperity, demand revives and cost again can be covered. It is not with the conditions incident to the business cycle that we are concerned here, however, but rather with dislocations of demand and supply growing out of overexpansion of fixed plant, major changes in demand, and advances in productive technique. When these are the causes of the inability of the producers to cover all their costs, the duration of the difficulty depends on (1) the durability of the fixed capital, (2) the trend of demand, whether stationary, rising, or falling, and (3) the extent to which producers are able to predict future changes in demand. For example, if the capital in the industry is very durable, and if a large part of it is new, the period of readjustment may be relatively long. It may take years for the producers to reduce the amount offered for sale to a point where price equals cost. Again, if the demand is increas-

ing, the period will be shorter under any given situation in respect to the durability of capital than it will be if demand is stationary or falling. Finally, if the producers misjudge the trend of demand and interpret each temporary revival of demand as a change in the long-time movement of demand and on that account replace old or build new plants, the readjustment of output to demand may be long in coming.

It may well be asked whether there is any truth in the statement that price tends to coincide with cost of production in industries requiring large amounts of fixed capital. Is it not more nearly correct to say that in these enterprises speculation plays so great a part that only prime cost affects price? Is it not true that the investment in specialized capital is so largely a gamble that no relation between supplementary cost and price can exist? To these questions two answers can be made. In the first place, if competition exists, the price of a good cannot remain for very long above the sum of its prime cost and supplementary cost, since competition both between businesses already established and between them and new concerns will result in such an increase of production that the gap between the cost and the price will be closed. In the second place, capital must all wear out eventually, and that which has proved unprofitable will hardly be replaced unless, as has been said, conditions seem to point to prices high enough to cover the going return on new investments. But, after all, too close an equality between cost and price must not be insisted upon, for this relationship is not one established by the forces of nature apart from the action of men themselves. The tendency for the values of freely mobile agents to agree with the values of the products to which they contribute is the result of many forces operating on the demand of consumers and on the entrepreneurs who control production. A perfect coincidence between the prices entrepreneurs have paid for the agents of production and the prices of the products cannot be expected as long as unforeseen forces are constantly interfering with the calculations of producers.

There are certain cases in which the coincidence is permanently interfered with. In the trade between two nations goods do not exchange for prices that yield the agents of production the same return in both countries. It may well happen that a unit of silk produced in Japan exchanges for a unit of steel produced in the United States (comparative money prices) at such a rate that labor in the steel mills in the United States can be paid higher wages than labor of comparable skill in the production of silk in Japan. The reason for this discrepancy is the fact that the Japanese workers have not the alternative of working in American steel

mills. There is even a certain amount of immobility of the workers between different sections of the United States. Consequently the same grade of labor may not receive the same wages in the production of different commodities coming from different sections. Finally, when a unique natural agent is required for the production of a single commodity, the principle of opportunity cost is not applicable. The normal price of the commodity into which this agent enters is, then, the sum of the prime and supplementary costs when the agents having alternative uses are applied marginally on the unique agent.

In recent years, in the United States, we have witnessed an important instance of the attempts of producers to adjust supply to the demand. Before 1914, agriculture in this country was in a fairly stable position. Prices were relatively high for grains, cotton, and meat, as compared with the prices of manufactured products. But there was no very marked tendency for the agricultural area to expand. Then came the World War. The demand for food and clothing increased greatly. The agricultural area was rapidly expanded, new techniques were developed that increased the productivity of land already cultivated, and more capital was applied to the land. But after peace had been made and the extraordinary demand of war had ceased, it was found that the agricultural areas, not only of this country but of other countries, were overdeveloped. Moreover, the new methods of production, such as improved strains of wheat and corn, better feeding methods, and new machines (among them the harvester-thresher), had made some of the new lands more profitable to cultivate than the old lands. Finally, there were many farmers who did not possess the necessary industry, judgment, and managerial ability to carry on an independent business under the new conditions. This was especially true in view of new techniques and the larger capital requirements of farming.

Under the type of adjustment described in the preceding sections of this chapter, the reduction of the supply of agricultural products would have been made by the reduction of output by the better farmers, by the abandonment of farms in unprofitable areas, and by the transfer of the incompetent farmers to other occupations. But the adjustment required was too great to be accomplished by these means without imposing great hardships on large numbers of people. The adjustment was slow in coming for a number of reasons:

1. During the war and even before the war an opinion had grown up that the government should interfere on behalf of the farmer. Particularly, special credit facilities had been provided for this group.

During the war the government urged the expansion of the farming area and the growing of certain crops. People had come to look to the government for direction and aid more than they did at the beginning of the century.

2. Production on the better farms and by the better farmers did not decline when the prices of agricultural products collapsed in 1920-1921, although we should have expected a decline from the preceding theoretical discussion. All the investment in land and buildings, as well as fences, drainage, and much agricultural machinery, represents sunk cost—investment in things that are durable and must be used according to the original intention of the investor or not at all. Moreover, farming is a “way of living” and not a job. The labor of the farm family usually must be employed to produce the commodities for which the farm is fitted or be used for nothing whatever. There is very little alternative use for land, buildings, machinery, and labor. Hence these investments do not give rise to variable but to overhead or constant costs. In terms of Fig. 15, the constant-cost curve, CC' , constitutes a greater part of the total-cost curve, TT' , than it would in merchandising or in many sorts of manufacturing.

3. Family farms are not operated on strictly business principles. The fact that the business enterprise is also the home of the farmer often keeps farm families on land when they should move to other land or to the city. Farming is not directly competitive as are most urban employments. It is more difficult for one farmer to copy the methods of another farmer than it is for one mechanic to copy the work of another. Unlike the merchant the farmer does not have held before him every day a test of his efficiency. When the merchant sees the buyers going to another store he is usually aware that he is being beaten in the race to attract custom. The merchant who loses patronage in any large measure is soon out of business. The farmer, who has relatively less variable expense of production, may hang on a long while after his business has become unprofitable. Often he finds that he would have done better had he sold out at the beginning of a price decline, invested in government bonds, and gone to work as a hired man. But the knowledge comes too late, and because men do not like to take orders after having been independent workers for some time many would not take these steps even if they could forecast a more profitable future.

4. The overcapacity of farms which resulted largely from war demand and which carried over into the decade of 1920-1930 came at a time when there were rapid technical improvements. In such a period

of transition from old to new methods, those who are financially unable to adopt the new methods or are unable to comprehend and apply them always suffer. There always appears to be an oversupply and an unduly low price for the goods of this group.

As a result of all these and other conditions that cannot be enumerated here, the demand for government interference became too insistent to be resisted. At first it was assumed that if the national government should buy the "surplus" of wheat and cotton, prices would rise and the stocks on hand could be disposed of without unduly depressing the price. But the prices of these two products fell even lower after 1929. Legislation enacted in 1933 established the Agricultural Adjustment Administration, which then offered contracts promising bonus payments to farmers who reduced their plantings of cotton, wheat, tobacco, and corn, as well as some less important crops. Partly as a result of the reduction in acreage thus achieved, partly as a result of unprecedented droughts, partly as a result of devaluation, and partly as a result of rising urban incomes, the prices of these agricultural products were greatly increased.

Farm income was increased, but nothing was accomplished by these restrictive measures to bring about a better adjustment of the productive capacity of the farms to the demand for products. Very little was done to remove farmers from unfavorable locations and still less to remove incompetent farmers from the business of farming. That is to say, the unfavorably situated producers are still in a position to produce, but not in a position to earn what is deemed an adequate income. The government is still (1937) contributing to their support. Certain other aspects of the government's program for the betterment of the incomes of farmers will be discussed in later chapters. Here it is sufficient to point out that the AAA did not aid materially in the adjustment of productive capacity to the demand for the products of agriculture.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Normal price is sometimes said to be the "right price"; it is also said to be the "just price" of a good. If price is above normal, in what sense is it not "right"? in what sense, if any, is it not "just"?
2. Cost of production governs price. A merchant who computes his costs correctly will never lose money. Is this statement correct?
3. Cost governs price only so far as it affects the amount of goods to be put upon the market in the future. Hence the fact that goods now offered for sale cost a certain amount has nothing to do with price. Is this statement correct? Is it partly correct?

4. It has been said that under normal conditions all cost is prime cost and that under market conditions nearly all cost is supplementary cost.

5. Can the curve VV' in Fig. 15 ever lie below the curve CC' ? Under what conditions?

6. It has been said that the interest on the mortgage on a farm is a cost of producing crops on that farm. Is this a correct statement? Suppose that the loan secured by the mortgage was used to buy land that has since declined in price? was used to buy electric refrigerators? was used to pay the expense of an illness?

7. Is a contribution of a business concern to a charitable enterprise an expense of production? Is the cost of insurance of buildings? Is the cost of bonding employees?

8. Make a cost diagram such as Fig. 15 on cross-section paper and show that when price rises above lowest total unit cost it will be more profitable for the producer to expand his output to the point where the price line intersects the marginal-cost curve than to produce only ON units.

9. It has been said that if the state assumes the responsibility for making good the losses that producers may sustain when price falls below the cost of production it must also be given the power to determine who may produce and how much may be produced by all producers. Explain.

CHAPTER XI · Interdependent Prices



THE PRINCIPAL TYPES OF INTERDEPENDENT PRICES

Thus far in our discussion of normal price we have given little attention to the fact that the prices of two or more finished goods or two or more producers' goods may be so related to each other that a change in the demand for or the supply price of one will affect the demand for or the supply price of the other. Simple conditions have been assumed for the purpose of making clear the fundamental relations of cost and price. A few complicating conditions of the type we are about to study, of course, have been taken into account in preceding chapters. It has been shown that one consumers' good can be substituted for another if they do not differ too much in quality, serviceability, and technical use. When such a substitution is possible, the price of either good cannot change very much, demand remaining the same, without some effect on the price of the other. We have seen also that because many different goods are derived from the same agents of production, the normal price of one good must be high enough to attract some of the agents away from others in the same group. But these are relatively simple cases of interdependence. In this chapter some instances of more complex interdependence will be discussed.

The phenomenon of substitution is not limited to consumers' goods. Raw materials of several sorts are frequently available for the manufacture of finished goods that go by the same name and serve the same uses. A table can be constructed from a number of different woods, all so nearly alike that only the expert can distinguish them. In other cases the substitution of a different raw material makes so great a change in the appearance or the use of the finished product that two different finished goods are produced. In these cases the one raw material cannot be put in place of the other unless the consumer is willing to submit to the substitution of a different finished good. Houses can be built of brick, stone, cement, or wood, but the use of each material results in a slightly different house, and a substitution of one finished good for another has occurred. Yet there is no question that the various

building materials do compete with each other and that they are frequently substituted for each other in house construction. When two or more producers' goods compete among themselves to supply the same consumers' good, or to supply slightly different consumers' goods that still can be regarded as substantially the equivalents of each other, they are said to constitute *rival supplies* for the finished good or goods in question.

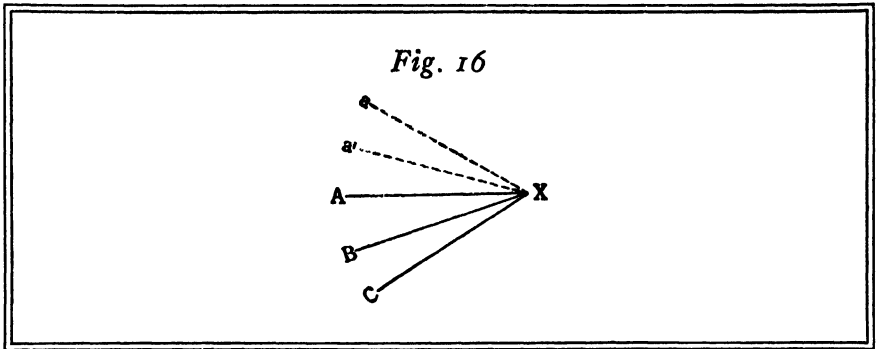
The converse of *rival supply* is *joint demand*. To build a house, a contractor requires not only the materials for framing and covering, but also hardware, glass, finishing lumber, and plaster, as well as several different kinds of labor. The demand of consumers for houses is a joint demand for all these requisites for house construction plus the services of the contractor. But the term "joint demand" must not be taken too broadly. The owner of the house doubtless will require more or less furniture to make it habitable, and he will also have food cooked and served in it. But the demand for housing is not a joint demand for building materials and food or for building materials and furniture. It is true likewise that many of the things we consume in the daily round of living are used together, but this fact does not make the demand for them joint.

Joint demand and rival supply have been contrasted because the first is a condition in which the demands for the several agents are so connected that when one is wanted the others likewise are wanted. Rival supply is a condition in which the supplies of the several agents are so connected that when one is wanted (for this particular use) the others are not wanted. The distinction is shown in the diagram on the opposite page, Fig. 16.

The demand for the finished good X is a joint demand for the agents A , B , and C . But it is not at the same time a demand for A , a' , and a , since the latter two are rivals of factor A . In other words, A , B , and C are complements; all are required in order that the good X may be completed. But A , a' , and a are competitors in supplying X ; and if the supply of A runs short, or if it goes too high in price, a' or a can be used in place of it.

When two or more different finished goods require the same factors of production, the demands of consumers for these finished goods are said to be *rival demands* for the agents used in producing them. Aluminum is used in making kitchen utensils and automobiles. Hence the demand for aluminum for making pots and pans is a rival of the demand for it in the making of automobiles. Common labor is used to

build railways and to build houses, and the demand for each is a rival of the demand for the other. Like joint demand, rival demand may be interpreted too broadly. All finished goods are rivals for the use of land, labor, capital, and entrepreneurship, and therefore are related in price. But this relationship is a very general one; and after specialization has been introduced, the factors divide into relatively noncompeting groups, which we have called agents of production. A true case of rivalry in demand occurs only when the same agents are wanted to produce two different finished goods. For example, the demand for



cheap factory-made furniture is not a rival demand for the better wood or for the highly skilled workmanship that goes into more expensive furniture.

When one productive process or one raw material always gives rise to two or more finished goods or two or more raw materials, we have a case of *joint supply*. The classical examples are cotton fiber and cotton seed, gas and coke, and the numerous products derived from the slaughter of cattle. Beef is the major product of the packing house, but in order to obtain it the farmer and the packer must produce also hides, glue, tallow, fertilizer, and other by-products.

RIVAL SUPPLY AND PRICE

When a finished good can be obtained from several raw materials, it may happen that only one of them will be used. Thus rails for American railroads are made of steel to the exclusion of other materials, although other kinds of iron, and wood faced with iron, have been used in the past and could be used again. Likewise copper is the only metal used in making wire for the transmission of electric current for long

distances, although other metals are possible. For technical reasons or because of expense these potential substitutes are not utilized. Hence it makes no difference what may be the price of iron-faced or iron rails; none will be substituted for steel as long as the price of steel remains near the present level. Likewise the price of platinum may rise or fall a great deal without affecting the use of copper for the transmission of electric current. At present platinum is far too costly to become a rival supply, and iron is too inefficient.

There are, however, many cases of rivalry in supply where the prices of the rivals are closely related to each other. Oak, birch, maple, and hard pine are used for flooring buildings. For some sorts of house no rivalry exists, and oak is preferred over all the others. But in still other types of construction maple or birch may be so nearly the equal of oak that the builder will choose whichever is for the moment the cheaper. On still other sorts of building the rivalry may be between hard pine and maple; and if one or the other rises in price, its rival will be used.

The conditions that determine whether goods will compete with each other are (1) technical requirements, (2) tastes and whims of the consumer, and (3) the prices at which the goods can be obtained in the market. The exclusive use of steel rails and of copper wire arises from the technical superiority of these two metals for their respective employments. The use of silk instead of wool or cotton in the manufacture of certain articles of apparel is determined in part by technical conditions but chiefly by the fact that the users of cloth prefer silk. But steel supplanted iron in the manufacture of rails only after the Bessemer process had greatly reduced the cost of steel, and the use of silk in place of other materials came about after the development of silk-growing and the improvement of textile machinery had made possible a reduction in the cost of silk cloth. Other woods have become competitors with oak for flooring in regions where the latter is not available and in all parts of the United States since the increasing scarcity of oak timber has greatly advanced its price.

Substitution affects prices in three ways: (1) When the demand for a good increases, the existence of a rival supply may prevent its price from rising as high as it would have risen had not the substitute been available. The existence of the substitute for a good causes its demand curve to be flatter than it would be otherwise. (2) Progress of the technique of production may bring in a competitive good in such large amounts that the price of the good for which it is a substitute will be

greatly reduced. The demand curve becomes less steep, and it moves nearer to the OX axis of the price-quantity diagram. (3) The demand for a good may be increased by the discovery of new uses for it—that is, when it becomes a substitute for other goods. The demand curve moves upward and to the right, and it also may change in slope.

The conditions of substitution may be classified roughly as follows: (1) complete substitution, (2) potential substitution, and (3) partial substitution. When substitution is complete and a new good entirely displaces the old one, there may be no price relation whatever between the two goods. Railroads are not likely at present to be displaced by ox carts, and the price at which the latter would sell, if manufactured, has nothing to do with the price of railway equipment. The existence of a potential substitute acts as a check upon the rise in price of the good in use; and whenever the divergence in price between a good and the price of its potential substitute becomes greater than the difference in their marginal demand prices, substitution will take place on a larger or smaller scale and the rise in the price of the good or the fall in the price of its substitute will be checked.

The effects of partial substitution are much more complicated than those of complete or potential substitution. The multiplicity of goods that satisfy the same types of wants gives the principle of substitution the greatest significance in modern consumption and consequently in the price-making process in the consumers' market. The most casual observation tells us that the desire for food, for clothing, for shelter, and for amusement and display may now be satisfied by a great variety of goods. To a considerable extent the existence of this variety results only in differences in consumption habits among individuals and among different social groups. No one person is likely to make use of all the possible means for satisfying his demand for food, clothing, and other ordinary wants; and if he takes up a new article, he is likely to drop the old one entirely. On the other hand, it is true that variety in individual consumption has been greatly increased, and the typical consumer now uses a vastly larger number of goods to gratify his wants than he did a hundred years ago.

A consumer often can make use of two or more goods for the satisfaction of the same want. If the two goods have for him exactly the same utility or desirability, that is, if the curves representing his personal demand for the two exactly coincide, then he will use the one that is lower in price. But such cases doubtless are not very common. Usually there is a preference for one or the other of the two. For

example, the owner of a motor car buys oil of a superior brand, A, instead of an inferior brand, B, although he might use the latter if it were cheaper than it is and if A remained at its present price. If we ask what will determine the point at which substitution of B for A will take place, the answer will be found to rest on the principle of marginal demand. Let us assume, for the sake of simplicity, that the owner uses one gallon of oil every month, or twelve gallons a year, no matter which brand is used. If the ratio of the marginal demand for A to that for B is as twelve to ten, he will buy the former as long as the ratio of the prices of the two is not more unfavorable to A than twelve to ten. If, however, he comes to believe that the better brand is something less than 20 per cent superior to the poorer, he will substitute the latter for the former. If the substitution takes place on a large scale, then the dearer (and superior) oil will be reduced (provided such reduction does not depress price below expenses of production) to the point where the two goods are again in competition.

Very often the two rival products are not capable of being put on the market in the same amounts. The output of the superior good may be limited by natural scarcity or by higher costs of production. Here we find the following adjustment of demand and supply: The more expensive product will be used by those for whom it has a relatively high utility for all purposes and by others for whom it has a high utility for special uses. In terms of the preceding example, if the superior oil is scarce, it will be used by those who believe it to have higher utility in proportion to its price than its rival, and by others who believe that for certain sorts of work it is superior to the cheaper oil. But it should be noted that the principle of diminishing utility operates to check the consumption of the A brand in the case of the latter type of user. If it is consumed in large quantities, its marginal utility declines, and finally a point is reached where it is a matter of indifference to the consumer whether he buys A at thirty cents a quart or B at twenty-five cents. If at this point the available output of the producers of both sorts of oil is sold, then thirty cents and twenty-five cents will be the normal competitive prices of the two lubricants.

To show the relation of the principle of substitution to the principle of cost of production, we may assume that the better oil cannot be produced at a cost of thirty cents. In that case the output will be gradually curtailed to the point where, marginal demand price having risen, the prices of the two rivals are proportional to their costs. If A costs thirty-five cents, then it will be withdrawn from those uses where its utility

is lowest (because of the refusal of consumers to pay that price for such uses) until a point is reached where its marginal demand is as thirty-five to twenty-five compared with the B brand. It might be hastily concluded from this statement that demand has nothing to do with the relative prices of the two goods, since each sells in the long run for a price approximately equal to its expense of production. But this conclusion is correct only on the condition that the two goods can be produced at constant cost, that is, that it costs the same amount of money per unit to produce a large quantity or a small quantity. If the costs of production vary as the outputs of the two brands are increased or decreased, then demand will control equally with cost of production.

JOINT DEMAND

When two or more goods are used to produce a finished commodity, the purchase price which the consumer pays for the article must be divided between the owners of these goods. The same idea may be expressed by saying that the costs of the several agents necessary for the production of the good must be covered by its price. The principles according to which this price is shared by the major factors of production—land, labor, and capital—involve the theory of the distribution of the annual income of the nation, which is the subject of later chapters. But the theory of distribution does not include the more limited relation of the agents of production within the same industry, and it does not include the price relations between different grades of materials or somewhat different types of land used in producing the same good. These more restricted relationships are a part of the study of price-making.

The theory of joint demand can be explained best by the use of concrete examples. We wish to know (1) what effect changes in the demand for the finished product will have upon the supply prices of the various factors or agents that are concerned in the production of the finished commodity. We wish to know (2) what will be the effect upon the supply prices of the various agents when the cost of one of these agents increases or decreases, and finally (3) we must inquire into the effect of an increase or decrease in the cost of one of the agents upon the price of the finished commodity. Let us take for our illustration the manufacture of a good, but not the most expensive, grade of furniture. We may assume that a given article, say a table, costs a hundred dollars to manufacture and deliver to the consumer. At that price two thou-

sand tables will be taken off the market. One hundred dollars is the normal price of the table. Its cost is made up as follows:

Raw materials	\$10
Unskilled factory labor	10
Skilled factory labor	20
Capital and management (factory)	20
Cost of distribution (dealers' margins)	40
Total cost	<u>\$100</u>

The supply price of the table thus is seen to be made up of the supply prices of the agents that are necessary to produce it. These separate supply prices, we may assume, are determined by the competition of other industries. The consumer, by his demand, determines how much shall be paid for 2000 units of the article, but he has nothing to do with the distribution of the total price among the factors.

Now suppose that the capital and management costs in the factory rise from \$20 to \$30, and the total cost to \$110. What will be the effect on the price at which tables will sell? What will be the effect on the prices that will be paid for the other agents of production? The demand for tables presumably has a certain degree of elasticity; and if the producers attempt to raise the selling price to \$110, sales will decline. Let it be assumed that they fall to 1800 units. But if production is cut to 1800, some part of the agents must lie idle, or they must shift to other industries. If none of them can shift, then competition will force them all—except capital and management—to take lower prices for their services, the total supply price will decline again to \$100, and 2000 units will be sold. On the other hand, if all the agents can change quickly without loss to other employments, the supply prices will remain as they were before the disturbance took place, except that capital and management will now cost \$30 instead of \$20, and the total normal supply price will be \$110. A stable equilibrium of demand and supply is arrived at when 1800 units are sold at that price.¹ We conclude, therefore, that if one of the agents jointly demanded to produce a consumers' good increases in cost, and if all the other agents can be transferred to other employments without loss, then the normal price of the good will rise by the full amount of this increase.

If the agents of production can move to other types of production, the consumer bears the increase in the cost of any one of them, and of

¹ Throughout the discussion it is assumed, for the sake of simplicity, that the cost of capital and management, per unit of output, remains the same (within limits) regardless of the volume produced.

all of them, if it happens that all rise in price. This is the long-time, or equilibrium, effect. But between the time when the cost of one agent rises and the time when the shifting of the agents has been completed price does not rise by the full amount. If the increase in cost is not foreseen by producers and preparation is not made for reducing output, price may remain as it is for a time. Then the agents begin to move, and price rises gradually to \$110.

Let us now modify the foregoing assumption by supposing that only one of the agents cannot move to another industry without very great loss. For definiteness, we shall assume that the skilled laborers cannot move unless they are willing to accept the wages of unskilled workers. Our problem is to explain how the increase in the cost of capital and management will be borne. As before, the producers will attempt to pass the added expense to the consumer; and, as before, the consumer will retaliate by refusing to buy as many as 2000 tables. There will be idleness in the industry, and the mobile agents—dealers, unskilled laborers, and the producers of raw material—will move to other industries. But the skilled laborers cannot move unless they are willing to take the wages of common labor. Some may be unemployed, or all may be employed only part of the time. If there is competition among them, their wages may fall to the point where they can do better by becoming common laborers in other lines. Wages will fall to this point unless, before it is reached, the sum of the supply prices of the various agents again becomes \$100. At that figure all the agents can be employed again and the total output can be sold for \$100 a unit. From the conditions laid down in this second assumption we may draw the second conclusion concerning joint demand. Under these conditions a rise in the price of one of the agents jointly demanded is borne by the immobile agent. Or it is borne by this agent until its price has fallen so low that it is more advantageous to shift it to another and lower use, in this case to employment at common labor.

In the preceding paragraphs we have examined the effects of a rise in the cost of one of the agents demanded jointly with others for the production of a finished good. If the change is the reverse, if the agent falls in price, the reasoning is the same. But in this case we find that the conclusion depends upon the mobility of agents *outside* the industry. If the capital and management costs decline to \$10, and if all the other four agents can be readily increased by additional supplies drawn from other industries, then the price of the table will fall to \$90. If, however, the supply of skilled labor cannot be increased, then the work-

men already in the trade will absorb the decline in the costs of the cheapened factor. Price will tend to remain at about \$100, and the skilled labor will be paid \$30 instead of \$20. There will, of course, be a period of transition and instability during which the price of the tables fluctuates around \$100.

A decline in the demand for tables at first affects adversely all the agents employed in their production; but after time has been allowed for the transfer of the agents to other uses, the consumers again pay the sum of the normal supply prices of these agents. If all can be transferred, normal price remains at \$100, but the volume of production is curtailed to the point where the demand absorbs it all at that price. If some of the agents are immobile, they must accept a decline in the price, and the consumers purchase tables at something less than the former price of \$100. Conversely, if the demand increases, and if there is complete freedom of movement of agents *into* the industry, the only result of the change in demand is an increased volume of production with sales at the same price as before the change took place. If, as we assumed above, one of the agents cannot be increased, that agent profits by the rise in demand for the finished commodity.

The supplies of the agents of production available for any industry are practically always elastic *if time is allowed for movement into and out of the industry*. Even skilled laborers who are narrowly specialized can be increased in numbers if time is taken to train new workmen. Their numbers can be decreased too if lowness of their wages keeps young men out of the trade. Likewise capital, even though it is highly specialized and durable, wears out in time and is not replaced if the returns to it are less in the furniture industry than in other lines of manufacturing. Hence the practical problem involves the changes in the prices of the agents during the period when the shifting is either impossible or is only going on slowly.

There are also other complicating conditions, most of which we have ignored in the preceding analysis. In the first place, some degree of combination is likely to be found among all groups of producers, for unlimited competition is nearly as rare as ironclad monopoly. The advantage in resisting a decline in price or in claiming the benefit of an increase in demand rests, other things being equal, with the group that can maintain the closest combination. A strong trade-union among the skilled workers in the furniture industry may prevent a reduction of wages when the capital and management costs increase, and it may enable them to prevent competition from outside the industry when the

demand for tables increases. It may readily accomplish these results if it is not opposed by a countercombination among factory-owners and if the owners have no other use to which the capital and management can be put. Some advantage may be retained by the union even if the owners are well organized, provided their organization is less effective than that of the workmen. But while the union may prevent a reduction of wages, it cannot prevent a reduction of the number of workmen employed. The owners cannot be compelled to employ their capital and management permanently at less than the going rates in industry at large.

A further complication arises from the fact that bargaining power plays a part in the determination of the supply prices of the agents jointly demanded. Shrewdness, knowledge of the market conditions, reserves of money to be used in fighting countercombinations, and many other forces have some effect in actual price-making. But neither side is likely to have a monopoly of these advantages for very long. And in the long run the normal forces of natural scarcity, real costs, and competitive alternatives will exert a controlling influence.

RIVAL DEMAND AND PRICE

The rivalry of consumers for the goods they use is largely a simple competition; that is, while there are many competing buyers who demand any one good, their demand for it is based chiefly upon the desire to use it for the same purpose. There is, for example, only one type of use for a hat or a pair of gloves. On the other hand, many producers' goods are wanted for numerous distinct uses, and the more elementary and unspecialized such goods are the wider is the variety of demand for them. Common labor and raw materials of many sorts are examples of goods that have literally a multitude of rival demands.

When a good has several uses, some are usually more important than others. The least important use to which a good can profitably be put is called its marginal use. Thus, if we should investigate the various uses of copper, we should find that for certain electrical machines a small quantity is practically indispensable. This, we shall assume, is its most important use. Then follow less important uses until a final, or least important, use (say, roofing houses) is reached. The marginal use is defined as that employment of any good which would be discontinued first if the price of the good should rise. When copper is five cents a pound, a considerable quantity may be used for roofing; but

if it should rise to a dollar, none would be employed for that purpose, and then a new marginal use would be established.

The price of any good for which there are rival demands tends to agree with its value in its marginal use. If copper is worth, in the opinion of the users, no more than five cents a pound for roofing, then it must sell for that figure in all the more important employments to which it can be put, for electrical machines, wiring, and the like. Now, because of the principle of diminishing returns, a good may have many marginal uses. Thus common labor may find its least important use in cleaning streets, in constructing railroads, and in making steel. In any industry there are many tasks that would be slighted if the wages of such labor rose appreciably. These are all marginal uses. *But they are not marginal in their entirety.* If wages rose, railroads would still be constructed, streets would still be cleaned, and steel would continue to be produced with the aid of much common labor. But some of the least important tasks, such as giving the right of ways of railroads a neat appearance, would be omitted; streets would be cleaned less thoroughly or less frequently; and various tasks in steel mills would either be dropped entirely or be performed less frequently and less thoroughly.

The value of any good in its marginal use sets the price for the entire market. It determines how much the market as a whole will be willing to pay for a given quantity of the good. If copper has been produced in such quantities that finally some of it must be used for roofing, then the price that builders stand ready to pay for it determines the price that other users will have to pay. On the other hand, it would be erroneous to suppose that the marginal use of any good solely determines its normal price. Whether copper will be used at all for roofing houses is determined by the quantity of it in the market and by the amounts other users are willing to take at prices that preclude its employment in place of other coverings. The supply of any good offered for a given use is fixed by the quantity produced and by the amounts demanded by users who are able to pay more than the prevailing price.

The theory of rival demands explains many price changes that at first sight appear inexplicable. The price of a good may advance without any apparent change having taken place in the conditions governing its total output. The wages of common labor used in making steel sometimes have risen simply because railroads have undertaken new building or extensive reconstruction and increased the demand for such labor and created a scarcity of it in the mill towns. During the summer

months many casual laborers leave the cities of the central part of the United States to work in the wheat fields. As a result laborers of this class who do not leave receive higher wages for city work. When the coming of winter sends these wheat-field workers back to the cities, urban wages for this class of work again decline. It follows that it is impossible to explain the price of any producers' good in terms of its total supply and the demand for it in a particular industry. Account must be taken of all the rival demands, and especial attention must be given to its marginal use or uses. Although it is true that the totality of demand for a good and the conditions governing its supply determine its price, it is at the margin that changes in demand and in costs are most effective.

JOINT SUPPLY AND JOINT COST

When two goods are the inseparable results of one productive process, it is impossible to ascertain the separate cost of either as far as that process is concerned. Thus cotton fiber and cottonseed, which are the joint products of the cultivation of the cotton plant, have no separable costs of growing. When two goods are supplied jointly, they of course must sell, in the long run, for a total price that will cover their joint costs. It follows that if the supply of one can be sold for only a relatively low price, the other must sell for a high price if the pair is to be produced at all. If the demand for one of these products either rises or falls, there will be a change in the price at which the other joint product can be sold. If, for example, the cotton from a given area of land can be sold for a hundred dollars and the seed for ten dollars, and if the sum of the two prices compensates the grower for all expenses, then a rise in the demand for cotton will cause the price of seed to fall. The increase in the demand will be certain to result in an augmentation of the cotton crop, which will increase the amount of seed as well as the amount of fiber. But the increased amount of seed cannot be sold at the same price as a smaller amount, its demand remaining the same, and on that account its price must decline. In like manner, a decline in the demand for fiber would raise the price of cotton seed. The only definite principle of costs that is applicable to joint products is that the sum of the selling prices of such products must cover their joint costs.

When two or more goods are produced jointly, it is impossible for the producer to cut off or to add one product, when he wishes to do so in order to increase his gross receipts, without at the same time affecting the amount of the other product. Such situations can proceed only

from natural circumstances over which man has little control. Many joint products, if not a majority of them, are therefore raw materials, like the fiber and seed of cotton, or half-manufactured goods, like grades of lumber. These commodities require further processing before they are ready for human consumption, and these processes entail further expenses of production. Hence the expenses incurred in producing consumers' goods are seldom joint in their entirety. Each of the products must be processed further after its emergence from the state of jointness in production, and each must bear its own separate expenses as well as such a share in the joint costs as will permit the producer to cover joint costs. When one of several joint products cannot be sold at any price, or when the costs of finishing it are above the price that consumers are willing to pay, it is not an economic good but merely waste. However, the discovery of new uses, the development of a new process, or the development of new demand may transform wastes into economic products. Examples are the by-products of the packing industry, cottonseed, and the smaller sizes of logs in the lumber industry, all of which at one time were considered valueless.

It is evident that the existence of joint supply and joint cost complicates the problem of normal value. Goods tend to sell for the sum of their costs of production; but if the products are supplied jointly, separate costs cannot be ascertained. Here demand plays the major role in fixing the price. If one of the two products is greatly desired by those who have the income to buy it, its price will be high, and its production may be pushed to the point where the supply of the by-products is so great that they become free goods; that is, they become free as far as the process of joint production is concerned, but not in respect to their separate costs. In railway transportation this phenomenon is met with when the traffic over a line moves predominantly in one direction. Since all cars moved in one direction must return in the opposite direction, it is economical for the road to seek return traffic at any rate above the difference between the cost of hauling a train of empty cars and a loaded train.

OVERHEAD COSTS NOT JOINT COSTS

Not all costs that appear to be joint are really so. In merchandising, for example, it is often not practicable to segregate the costs of selling different kinds of goods. Specifically, all advertising carried by the store may have some effect upon the sales of all departments and of all

goods. If it is successful, more people pass through the store; and although many articles are never mentioned in the advertisement, they benefit to some extent because more people are brought to the place where they are offered for sale, and they may justly be said to occasion some advertising expense. Yet it would be an impossible task to apportion the expense among them all. It would be impossible also to discover the proportions of the salaries of managers that should be charged to each type of commodity, although sometimes an arbitrary assignment is made. Such costs as the salaries of managers are usually carried into an account labeled "overhead" to indicate that it cannot be definitely apportioned.

Such overhead costs are similar to joint costs, and they affect business decisions in somewhat the same way. As long as the total sales are high enough to cover all costs, both those which can be apportioned and those which cannot, the store continues in business. But there are striking differences between overhead and true joint costs. When the demand for one of a pair of true joint products rises sharply and its production is increased, the other member of the pair also must be produced more freely. And, other things being equal, it will fall in price. This is the principal test of true jointness. When it is applied to the goods sold in a store, we find that a demand for one commodity does not result in an increase in the supply of other commodities. On the contrary, if the demand for one group of articles rises, the store may even contract the space and selling force formerly used for the other groups and use them to market the more profitable goods.

Still another test can be applied to determine whether the costs of two or more products are really joint costs. If a railroad raises the rates on commodities moving in one direction, the amount of traffic will tend to decline. The number of loaded cars moved in that direction having been reduced, the number moving back unloaded also will be reduced, and the necessity of offering unusually low rates on the back haul will be less. But if a store finds it necessary to raise prices in one department, that fact will not reduce the amounts of goods offered in other departments or enable the concern to ask higher prices for them. Again, if the separate costs of one of a pair of true joint products are increased, its marginal-supply price will rise, sales will decline, and less will be produced. The output of the other member of the pair will also be reduced, and it will rise in price. But if the special merchandising cost of one article increases and its price rises, the resulting fall in sales will neither reduce the supplies of other goods nor raise their prices.

SUMMARY OF THE THEORY OF NORMAL VALUE

Before proceeding to the next section of the theory of value and price, which is concerned with price-making under conditions of monopoly and imperfect competition, we may sum up briefly the principles of normal price thus far explained.

Under competitive conditions a good tends to sell for a price equal to the sum of its costs of production. Its costs of production are determined in part by the competitive demands of consumers for the various consumers' goods for the creation of which the agents of production are used. A factor, in the long run, must earn as much when used to produce one good as it does in the production of other goods. This gives rise to opportunity cost. But some consumers' goods are more highly prized than others, and it might appear that the agents on that account would earn more when applied to such production than when applied to producing other things. The solution of this difficulty is that such highly prized goods will be produced in larger amounts than commodities of little significance to the consumer, and in consequence their marginal-demand price will fall to the point where the last units of the factors used in making them will have the same imputed value as the last units used in making other things. The factors thus earn the same in all competitive uses.

The tendency of price to agree with cost of production must not be interpreted as meaning that whatever costs have been incurred by the producers will necessarily be covered by the prices offered by consumers for finished commodities. This agreement will take place only if the entrepreneurs in that group have forecast demand and supply accurately. When errors in prediction take place in respect to capital investments that are short-lived, the readjustment is made quickly. These errors in prediction give rise to the fluctuations of market price about normal price. But when too much capital has been invested in specialized, durable equipment, the adjustment may be so difficult to make that for long periods of time only the prime costs and part of the supplementary costs arising from the investment in specialized, durable equipment can be recovered by the entrepreneurs in the marginal group. Since there are no methods available at present for forecasting the conditions of demand and supply in the distant future, any investment in very durable capital is more or less of a gamble. This kind of capital may become for considerable periods a unique good, similar to special kinds of land, whose value is determined, not by its cost, but by the

demand for the finished product and the costs of the other agents demanded jointly with it to make the finished product.

The existence of joint demand, rival supply, and rival demand does not require any modification of the doctrine that price tends to agree with cost of production under conditions of perfect competition and complete mobility of the agents of production. But the condition of joint supply does require such a modification. When costs are actually joint, the combined prices of the joint products must be high enough to cover the joint cost plus the separate expenses of each; but no allocation of the joint cost to each product is possible, and therefore the price of a joint product is not governed exactly by cost.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. All the goods in the world constitute rival supplies for the satisfaction of the consumer's wants. Hence the prices of all goods are related to each other. Why, or why not?

2. If the demand for beef should increase greatly, how might the cost of shoes be affected? Why?

3. There are several different materials for roofing dwelling houses, some of which cost about the same, while others are much more expensive. How do you account for this? Why is each material not sold at a price considerably different from the others?

4. It is sometimes said that the transportation of freight and passengers is an example of joint supply. Is this correct? Why, or why not?

5. How, if at all, are the prices of automobiles, gasoline, and cement related?

6. A mechanic who earns six dollars a day at his trade when employed carves toy boats, on days when he has no work, at the rate of two a day. Is it correct to say that the labor cost of the boats is three dollars apiece?

CHAPTER XII · Equilibrium Price and Monopoly¹

COMPLETE MONOPOLY AND COMPETITION CONTRASTED

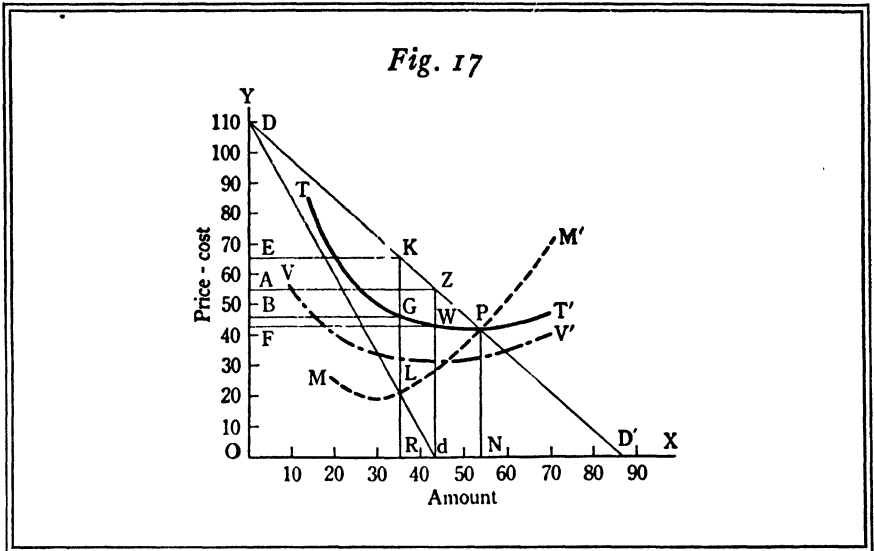
The essential difference between complete monopoly and competition is that when monopoly exists, there is only one seller, and when competition prevails, there are many sellers. If there are many sellers and only one buyer, the term "monopsony" or "buyer's monopoly" should be used. The fact that monopoly implies only one seller does not mean that there are no substitutes for the commodity sold by the monopolist. There are substitutes for nearly all commodities offered for sale in the market, and, as has already been explained, the existence of substitutes is one of the factors determining the slope of the demand curve.

In Chapter X it was explained that the seller in a purely competitive market is virtually unaware that a down-sloping demand curve exists. For example, a large wheat-grower, who sells annually 10,000 bushels, finds that it makes no difference in the Chicago price when he sells his crop. In fact, so small is his contribution to the national crop of 700,000,000 bushels that the price would be unaffected if he refused to sell at all. Conversely, he might increase production to 20,000 bushels without causing the average price for the year to be appreciably different from what it otherwise would have been. As the individual seller sees the demand curve, it has virtually infinite elasticity. It has been represented, therefore, by a straight line parallel to the quantity axis, OX , in Fig. 15 on page 158.

With these facts in mind we now contrast the position of the seller under pure competition with the position of the monopolist. The monopolist has complete control of the supply of the commodity he sells. The demand curves for all commodities are negatively inclined, that is, they slope downward and toward the right. Hence if the monopolist offers a large amount he must sell that amount at a relatively low price, and if he offers a small amount he can command a higher price. The

¹Part of the explanation in this chapter was first developed by Professor Edward Chamberlin in *The Theory of Monopolistic Competition*, 1933. See also Joan Robinson, *Economics of Imperfect Competition*, 1933.

monopolist not only will be aware of the inverse relation of demand to price, but will take account of this fact in determining the amount of the monopolized good to put on the market. In Fig. 17 the following conditions are represented: (1) a demand curve for a monopolized commodity, DD' , also called the general-revenue curve; (2) a marginal-revenue curve, Dd , which was explained in Chapter VIII; (3) the three cost curves, VV' , TT' , and MM' . The purpose of the figure is to show how monopoly price is determined under ideal conditions. These con-



ditions are as follows: (1) It is assumed that the monopolist knows the position of all five curves, or that he knows enough about the various types of cost and about demand to be able to forecast with substantial accuracy the prices that will be paid for different amounts of his commodity and what different amounts of that commodity will cost, not only in terms of average total cost but also in terms of variable and marginal cost. (2) It is assumed that the monopolist will not be interfered with by public authorities, no matter what prices he charges; otherwise he would be governed in setting his price by something other than economic considerations. (3) It is assumed that the monopolist is so secure in his monopoly power that the price he sets will not be influenced by the fear that a competitor may appear if price is set too high and profits are too large. Stated briefly, the objective of the monopolist is to secure the largest monopoly profit, that is, the largest gain above his average cost of pro-

duction. The price that will produce this maximum is not, of course, the price that shows the greatest excess of price over cost, because at that price only a small volume of goods could be sold, and although the profit on each unit would be maximized, the total profit would be small. Since the added revenue increases with each increase in sales until Dd intersects OX at d , it might be supposed that Od units would be offered in the market at a price of dZ or OA . This would be the answer if it were assumed, as in Chapter VIII, that no cost existed. But in this case it has been assumed that cost does exist. The monopolist, therefore, should continue to enlarge his sales from zero until the added revenue from an added unit of sales exactly equals the added cost of that unit. Since an amount, OR , can be sold at this point, the price he should charge under ideal conditions is RK , and the monopoly profit is $EKGB$. If more than OR units were sold, the *added* units would cost more than the revenue to be obtained from them, as is shown by the fact that the marginal-revenue curve, Dd , lies below the marginal-cost curve to the right of the line $RLGK$. It might be supposed also that the monopolist would produce (say) Od units or more, because at that point in the supply his average total cost, dW , is less than the price, dZ , and a profit could still be made on the sale of additional units. But this conclusion is incorrect. If more than OR units are sold, the price of the entire output falls, because the demand curve with which a monopolist deals is not the horizontal price line that confronts a seller in a competitive market. Additional units cannot be sold by a monopolist without a reduction in price. If OR units are sold, the monopoly profit is greater on each unit although the number of units disposed of is less. In the figure the monopoly profit when OR units are sold exceeds that profit when Od units are sold by approximately 20 per cent.

The figure, of course, is merely a geometrical representation of the economic facts. The principle may be stated without reference to the figure. A monopolist should enlarge his sales from zero to the point where an added unit of sales increases his gross receipts less than the added unit increases his total costs.

In the figure, DD' intersects TT' at P . It may be asked why the monopolist did not build his plant small enough so that OR units, rather than ON units, could have been produced at lowest total unit cost. The answer is that in this case it is assumed that if he had tried to build a plant with a capacity to produce OR units the point of lowest total unit cost would have been much above NP and might even be above RG , and monopoly profits would have been smaller. Of course it might well

be the case that the amount the monopolist could most efficiently produce would be OR units, and at the same time the demand curve might retain the position as shown in Fig. 17. If such had been the case, monopoly profits would have been greater. But it might be true also that the demand curve would intersect TT' much farther to the left than the point P , and the monopolist might still be unable to build a more economical plant than one that would produce not less than ON units at lowest total unit cost of NP . Obviously a multitude of combinations of the curves could be imagined, and each combination would produce a different monopoly price and a different monopoly profit.

CAUSES OF MONOPOLY

Conditions that give rise to a complete monopoly that can adjust prices in the manner explained in the preceding section are not common. They may exist under four conditions: First, one concern may own the sole source of the supply of one or more of the raw materials that go to make a finished good and therefore may have a monopoly of the production of that good. But the finished good must be nondurable if the monopoly is to be completely effective. For many years the DeBeers concern supposedly had a virtual monopoly of the supply of uncut diamonds from mines. But diamonds are extremely durable, and as soon as the price was raised too high cut diamonds in the possession of dealers and consumers came into the market.

Second, one concern may be granted the exclusive right to produce or sell a commodity or service in a certain market. This is common in the case of public utilities, such as the supply of water or gas or electricity for public use, street transportation, and the telephone. But these monopolies are subject to regulation by the public authority, and the management is not free to charge the prices that will yield the maximum profit. The state usually steps in to regulate prices, and it is almost sure to do so sooner or later if full advantage is taken of the fact that a monopoly exists. So great is the fear most utility companies have of public opinion that a majority maintain a public-relations department whose function is to watch the currents of such opinion and to combat by means of personal contacts and publications any attempt on the part of any group or individual to create an opinion that rates are too high or service inadequate.

Third, the state may create a monopoly through the grant of exclusive right to produce some mechanical contrivance. Such a grant is

called a "patent," and it protects the inventor or the owner of the invention against the duplication and sale of the patented article. Such patents, however, are not perpetual, and in the United States the exclusive right expires seventeen years after the date of the grant. A similar exclusive right is granted to the writer of a book. This is called a copyright. Here again the right expires after a time, but it may be renewed so that the protection extends for fifty-six years.

Fourth, monopolies may also arise out of the economies of large-scale production. Such monopolies are usually called natural monopolies. Many years ago it was said that no steel plant could be operated efficiently unless it had a capacity represented by an investment of \$25,000,000. Today the necessary investment would be much larger. The fact that so large an investment was required to build an efficient plant was supposed to keep others from attempting to interfere with the dominance of the United States Steel Corporation. It is doubtful if this was the true explanation of the position of that corporation, but the illustration is helpful. If the market for a good is so small that only one concern can possibly sell enough of the good to enable it to produce at lowest total unit cost, then the conditions that create a natural monopoly exist. If there are no economies of large-scale production, then it is improbable that only one concern will be able to find a place in the industry. There are a number of efficient steel plants in the country. Often, of course, rare goods and new commodities that have not yet secured a wide market may be monopolized temporarily.

Most of the so-called monopolies that exist in the United States at the present time, with the exception of public utilities and patent rights, are really partial monopolies. That is, they consist of several sellers who are able for various reasons to maintain prices above the competitive level but who are not able to raise prices to the level attainable by complete monopoly.

MONOPOLY AND PRICE DISCRIMINATION

The preceding section has shown the conditions governing the price a monopolist would set upon the good he sells. But in that section it was assumed that the monopolist was unable to charge different prices for different portions of his output. Often a seller cannot discriminate between buyers. If certain conditions are present he can. It would be futile to try to charge different buyers considerably different prices for cigarettes, because those who bought at a low price

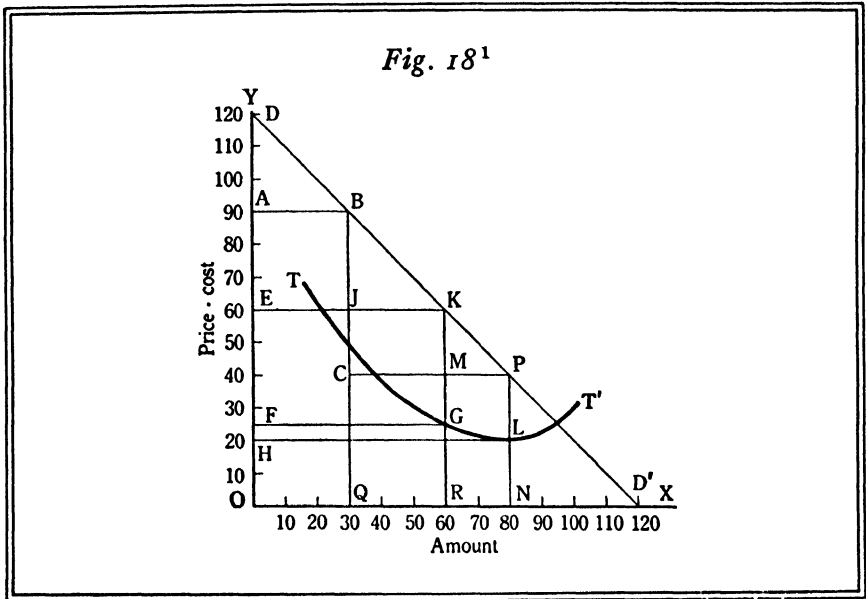
could and would sell to those to whom the seller attempted to sell at a higher price. But it is not impossible to sell to well-to-do people in a fashionable shop at a higher price the same quality of goods as is sold to people of lower incomes in more ordinary shops. And it is common for railroads to charge more for carrying a carload of machinery than for carrying a carload of coal the same distance. The shipper of coal cannot turn the service he buys over to the shipper of machinery.

Sometimes the producer can sell consumers different amounts of the same commodity at different prices. Usually he does this by selling a small amount at a high price and a larger amount at a lower price. Such prices are found under both monopolistic and competitive conditions. Under monopoly we find both gas and electric rates customarily higher for small amounts than for large amounts of consumption by consumers similarly situated. Under competition the buyer in large amounts receives a quantity discount. In both these illustrations differences in the cost of service justify some discrimination. Here we are concerned, not with the justification of the discrimination, but with the possibility of its being put into practice.

Ideally the profit-seeking entrepreneur who is a monopolist would obtain the greatest possible monopoly revenue if he could compel the buyers to pay for each unit of the supply the exact demand price for that unit. This is complete discrimination. In Fig. 18 this would mean that the monopolist would sell (say) 80 units and charge prices ranging downward from \$120 to \$40 a unit. His monopoly revenue would then be measured by the area *DPLH*. But he can seldom do this, because (1) one consumer may be able to resell to or buy from other consumers; (2) the monopolist may not be able to determine which consumers are prepared to pay high and which low prices; (3) the monopolist may not be able to compel the consumer to pay a high price for the units most urgently needed and then sell him additional units devoted to the satisfaction of less urgent wants at lower prices.

In Fig. 18 one type of discriminatory monopoly price is illustrated. Here a monopolist that could not employ discrimination would sell 60 units for \$60 a unit, or a total of \$3600. The average cost of this amount is \$25 (*RG*) per unit, the monopoly revenue is \$35 (*GK*) per unit, and the total profit is \$2100 (*EKGF*). Now suppose that the monopolist can divide the total market into two groups, one of which will pay \$90 a unit for 30 units and the other \$40 a unit for 50 additional units. It should be noted that if the monopolist can produce and sell 80 units (30 + 50) his average cost will fall to \$20 a unit. On the first 30 units he therefore

obtains a monopoly revenue of \$70 a unit, or \$2100, and on the next 50 units \$20 a unit, or \$1000. His total monopoly revenue is now \$3100 instead of \$2100, which would be his maximum if a non-discriminating price were charged. The monopoly revenue has increased from *EKGF* to *ABCPLH*. Of course, he has lost revenue represented by the area *JKMC*, but he has gained *ABJE* and the irregular area *FGMPLH*.



Some of the more important effects of such price discrimination are as follows: First, discrimination permits the monopolist to exploit the advantages of his monopoly more fully than he could otherwise. Second, the charging of a higher price to those whose demand is more urgent permits a more extensive scale of operations and thus permits the sale of a larger number of units, some of which are sold at a lower price. In the figure we find that instead of 60 units sold at \$60 a unit, 30 of these units are now sold at \$90 and 30 at \$40. In addition, 20 units are sold also at \$40 a unit. Under nondiscriminating price the average cost to the consumer was \$60 a unit, but with discriminating price the average cost is \$58.75. It must not be forgotten, however, that if competition had prevailed, or if the industry had been regulated by government authority, the average price might have been still lower.

¹ *DD'* is the demand curve and *TT'* the total-cost curve.

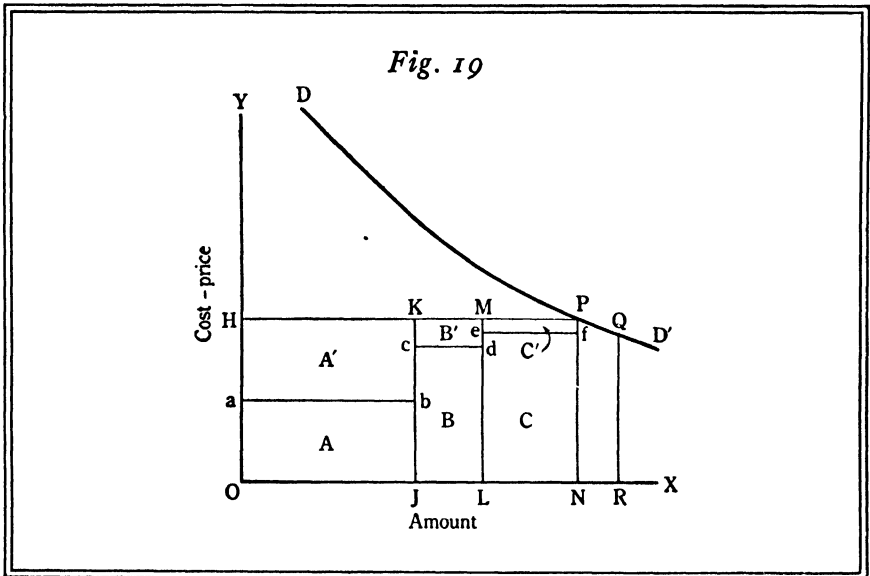
Different prices are often charged different consumers in the same market, but these prices are not necessarily discriminating; nor do they necessarily arise out of monopoly conditions. The consumer must often pay more per unit for small purchases than for large purchases. Consumers of electric current who have about the same type and size of house will usually pay different average unit prices for current if the amount of consumption is different. The explanation of these differences is that every purchase at retail involves some fixed costs, such as making out the sales slip, making change, and wrapping. In the case of electric current it has been found that it costs about as much to read the meter and to bill and collect, whether the customer uses 50 or 1000 units of current each month.

MONOPOLISTIC COMPETITION

When a few sellers (say from two to ten) compete in selling a commodity, a condition of monopolistic competition often exists. It is assumed for the purposes of this section that the sellers do not attempt to influence demand. Three different types of such competition will be explained. Each leads to a different result.

One type of such competition is found where one large concern and a number of smaller concerns sell in the same market. For example, the steel industry is dominated by the United States Steel Corporation. In some parts of the country one of the largest oil companies is the leader among a pack of several lesser competitors. The operations of these concerns, of course, are not made public. But if one may judge from the facts that do come to light, the leader in such a monopolistic competition sets the price, and the smaller firms follow. Sometimes the smaller firms are allowed to charge somewhat lower prices than the leader because otherwise they could not attract enough business to enable them to continue. Usually the costs of the smaller firms are higher than the costs of the leader or leaders. When that is the case the position of the various firms can be illustrated by Fig. 19. The demand curve DD' exists for the industry as a whole, but each segment of the sellers caters to a special demand curve. Thus, for example, there are millions of owners of motor cars who believe that the types of gasoline sold by one large seller are superior to those sold by other sellers. This means that they are willing to pay somewhat higher prices for the product of that concern than for the products of other concerns. Usually that concern does not charge higher prices but uses the

preference that exists to enable it to maintain a larger volume of sales than it otherwise could maintain. The broken line $abcdef$ represents the total average costs of the different segments of the supply. The rectangles A , B , and C represent the total costs of the three different groups of sellers. The supply OJ is put on the market by the leader, JL by the next most advantageously situated seller or sellers, and LN by the least advantageously situated group. The price is set by the leader at OH or NP .



The profit made by the leader is A' , by the next group B' , and by the last group C' . Of course the leader might set the price somewhat lower, say at Nf or RQ . Then more producers would be included in Group C, and less profit would be obtained by all groups, the leader included. Or, again, he might set the price at Jc , drive out all the producers in Group C, and expand his own supply. Historical instances show that the dominant firm in an industry has pursued all three types of price policy. Suppose that the leader decides to drive all the other firms out of business. First he must cut price below the cost to Group B (Jc) and keep it there until all the members of Groups B and C are bankrupt or are willing to sell their plants to the leader. Then the leader will have the entire field to himself and can charge a pure monopoly price, such as was explained in Fig. 17. But there are a number of reasons why the leader may not attempt to create an exclu-

sive monopoly at the present time in the United States. (1) He would probably arouse public opinion against himself, and this might lead to prosecution under existing antitrust laws or to new regulatory legislation. (2) He could not be certain that he would be able to prevent other firms from intruding into his monopoly after he had created it, especially if the profits were very attractive. For some years the Ford Motor Company was virtually alone in the production of a successful low-priced motor car. But its gains were so great that after some time the General Motors Corporation entered into competition. Today the latter corporation and the Chrysler Company divide the field with the Ford Company, and the profits of the first producer are not so great as they were ten to fifteen years ago. (3) The leader would incur losses while engaged in the price war to drive out competitors. That is, he would probably be unable to make the profits represented by A' in Fig. 19, and he might even find it necessary to sell below Jb , his average cost of production. Furthermore, if the monopolist expanded his output beyond OJ to prevent the intrusion of outsiders, his average costs might become higher and his profits no greater than if he produced only OJ units.

A second type of monopolistic competition appears when two or more producers compete on more nearly equal terms than those represented in Fig. 19. Sometimes these concerns enter into a combination and divide the field. That is, they charge the full monopoly price and divide the volume of sales that will produce that price among themselves according to some formula that is usually arrived at by the methods of diplomacy—threats, special concessions, and shrewd dealing. But if the conditions are favorable for monopoly, and if two or more competitors cut prices, all may be ruined. Then a reorganization of all the firms into one firm may take place, and for a time a monopoly price may be charged.

A third and more frequent type than either of the preceding appears when a limited number of competitors operate in sheltered markets where competition can enter only if the price is set too high. For example, there may be a brick factory in a city of 25,000 population, and it may supply the building needs of that city and of near-by smaller communities. But there is not enough business for two or more concerns to operate efficiently in this location, and unless a potential competitor believes that he can operate at greater efficiency than the existing producer and therefore can drive him out of business no local competition will appear. Other concerns at a distance, each having its

own local market, may compete. But if we assume that their product is not superior, they can compete only if they absorb the cost of transportation and other handling charges incident to delivering the brick in the home city of the local producer. It is evident that this situation is much the same as that illustrated in Fig. 19. But in this case each seller has a protected market, which others can intrude into only by adding something to the costs of production. Each producer will charge in his home city the price at which his distant competitor is willing to sell. And this price should be higher than the cost of production of the local producer. Competition exists but, owing to the cost of transportation, the competitors are not on the same footing, and the local factory may make considerable profit by reason of the partial monopoly that arises out of transportation costs.

COMPETITION AND MANIPULATION OF DEMAND

Whenever a seller or manufacturer can acquire a sheltered market into which other sellers cannot intrude without incurring additional costs, he will usually try to do so. To a limited extent every seller attempts to attach a clientele to his business. He does this by furnishing certain services not easily imitated by others, by personal contacts, if possible, and sometimes by paying more careful attention to the quality, assortment, and reliability of certain types of merchandise. Advertising, however, is the most powerful method of building a clientele in the United States. The devices used by advertising experts to work upon the emotions, prejudices, fears, and vanities of the buying public are numerous. So complicated have these devices become that the construction of advertisements is a highly skilled and difficult profession. But the objective of advertising is always the same, however much the methods may differ. The objective of the advertiser is to segregate a portion of the buying public and attach its allegiance to his wares. This means persuading a part of the consumers to buy without making comparison with the wares of others, or persuading them that the advertiser's wares are somehow preferable to the wares of his competitors.

If the advertising is successful the "regular customers" of a firm sometimes will buy without inquiring about the price charged by others, and sometimes they will inquire but retain the opinion that the difference in price is more than offset by the quality of the goods offered by the advertiser. Often, of course, the assertion of the advertiser is that he sells goods that are the equal of those sold elsewhere but at a

lower price. We are not here concerned with the social desirability of advertising but with its effect upon competition. When any considerable portion of the buyers in a market come to have a predisposition to buy one brand or to buy from one seller rather than another for some other reason than lower prices, the demand curve for a given seller's product ceases to be a horizontal price line. It takes on a slope, just as does the demand curve for a monopolized good.

The advertiser often can secure higher prices for his product than can be obtained for unadvertised goods. That is, he can obtain such prices from some consumers. Others are more sophisticated or more expert in resisting the direct and indirect emotional appeals. They cannot be persuaded to increase their demand price or to become regular customers by bold assertions of superiority or by a picture that suggests that unless they use a certain product they may miss some desirable experience, lose caste socially, lose their jobs, lose money, become ill, or suffer some painful accident. Those who are most affected by the advertiser's appeals are willing to pay the highest prices; those who are unaffected will not pay any price above that asked by other sellers for substitute goods. Between these two extremes is a more numerous group that will pay somewhat higher prices.

The demand curve that the advertiser thus creates for his particular product is not different from the demand curve representing the entire market demand for any type of commodity. Thus there is a distribution of potential buyers for any common article like toothpaste. Some will pay very high prices rather than go without; some will pay moderate prices; and some must be offered the good at low prices if they are to buy at all. This distribution, as has been explained in Chapter IX, arises from all the conditions that affect choices. If toothpaste were sold as nails, window glass, and wheat are sold, no seller could affect the demand curve and the price. But the advertiser segregates a portion of the market demand for toothpaste and establishes a differentiation in consumer preferences between his brand of toothpaste, other brands, and dentifrices of no brand.

In Fig. 20, TT' represents the total costs of a producer of an article for which no clientele has been established. Since competition is assumed, the price is set at NP by the total market competition, and the producer is just able to sell his output of ON units at cost. But he now resorts to advertising or some other kind of sales promotion and tips the price line, ee' , to the position of DD' , where it becomes a general revenue curve. In so doing, however, he incurs additional costs, and the total

It is often argued by those who defend advertising against the statement that it is mostly acquisitive and sometimes antisocial, that all sales-promotion work increases the volume of sales, enables the seller to reduce costs, owing to larger volume, and therefore costs the consumer nothing. Whether the enormous amounts of money used in pushing goods onto the market and in influencing the choices of consumers result in any benefit to society, that is, to the mass of consumers, cannot be answered categorically. The arguments usually advanced to show that such is the outcome can be examined only briefly. It should be noted first that only the most naïve will believe that the primary objective of advertising and kindred clientele-building activities is social betterment. The objective is usually to cut off the consumers from the competition of rivals. It is said, however, that there are many commodities for which the demand would be so small in the absence of sales promotion that even one concern could not secure a sufficient volume to enable it to secure the normal economies of machine production. This argument is doubtless good in a limited number of cases, but it obviously does not apply to thousands of advertised and branded articles such as automobiles, flour, shoes, clothing, breakfast food, and cured meats. Volume, however, does become a problem to the individual producer when a rival declares war and seeks to annex all the customers who would naturally fall to that producer. He must fight back or lose his business. Both must use sales promotion, and in the end the consumer pays for being cajoled into buying goods that he would have bought anyhow. Often he is induced to discard an article that is still serviceable and replace it with a new model. This style element, formerly limited primarily to clothing, has now appeared in automobiles, radios, houses, letter paper, and hundreds of other things. In many cases, of course, the innovations are real improvements.

Once a producer has built a large plant, having acquired a sufficient clientele to make full utilization of the plant possible, he is caught in a trap. If he does not maintain volume, the overhead costs rise rapidly. Even the variable costs may rise. Therefore, *from his individual point of view*, sales promotion is necessary to permit him to sell enough to keep costs down. The costs of sales promotion are just as *necessary* from his point of view as the costs of labor and raw materials. This is probably what those who argue that advertising reduces costs usually have in mind.

Partial monopolies are the result of these dyke-building activities of the advertisers, salesmen, and designers of "new models." They create

sheltered markets by working on consumers' preferences, which are only less effective in the protection of high profits than such natural obstacles as transportation costs and the economies of large-scale production. A further effect may be noted. Under pure competition the seller who cannot secure enough of the factors of production (who has inadequate capital) to make use of improved methods is quickly eliminated. But with competition impeded he may be able to charge prices high enough to cover his abnormally high costs.

In his recent work Chamberlin¹ has shown that when, through sales promotion, product differentiation, and other similar devices, the market comes to be parceled out among a group of rivals in the manner described above, it is unlikely that any of the rivals will be able to secure enough volume to enable him to produce at the point of lowest total unit cost. In terms of Fig. 20, the demand curve, DD' , cannot become tangent to the cost curve TT' or AA' at P or P' , so long as it has slope. Only a horizontal price line such as ee' can be tangent to either curve at its lowest point. The economic analysis of what has just been stated in terms of geometry is as follows: As long as many customers prefer one seller's wares to those of another, so long will the seller be able to reduce output and confine his sales to those who are willing to pay higher prices. When this condition has persisted for a time, the sellers may become so numerous that none can sell a sufficient output to enable him to reach lowest total unit cost, even though he should offer to sell at a price equal to that cost. There is not enough volume of demand to go around. By contrast, pure competition would force price down to the lowest total unit cost of the typical sellers. It would do this by eliminating some of the sellers—by assumption the least efficient. It therefore appears that the building of clienteles is one of the causes of chronic overcapacity in certain lines of business, chiefly those where manipulation of demand is most easily accomplished. Obviously, however, this cannot be the explanation of overcapacity in such businesses as agriculture, where manipulation is absent.

It is frequently said that if a seller can induce people to believe that a good is worth more than the sum of its production and distribution costs he has "created value." It is said that no matter how a consumer comes to have a demand for a good, even by the grossest misrepresentation, that demand is just as valid for the purposes of price analysis as a demand based on the most obvious biological need. With this state-

¹ E. Chamberlin, *Theory of Monopolistic Competition*. Cambridge, 1936.

ment we must agree. All demand rests in part on social influences as well as biological need. Here it is not the intention to argue the point whether advertising is socially desirable or undesirable. Its effect upon price is, however, our concern.

Because advertising and other sales promotion may segregate different portions of the demand and attach them to different producers, the entire market for any commodity comes to be broken up into a number of partially separate but yet connected markets. Advertiser competes with advertiser, and first the demand for one brand and then the demand for another rises and falls. The type of competition found in the sale of grains and of common unadvertised articles becomes tinged with elements of monopoly. Because of the variability of demand and the succession of new products, often different only in appearance, stability of output by a given firm becomes improbable. The reader is familiar with attempts of the producers of low-priced motor cars in recent years to catch the public's fancy first with one "improvement" and then another. Some of these innovations turned out to be not very useful, for example, free wheeling. Others find permanent acceptance, but many are of no consequence either for the efficient operation of the cars or for the safety and comfort of the users. When, because of a successful raid by one seller on the demand curve of his rival, that rival's sales decline, the latter finds himself with unused capacity. In many cases he cannot cut prices and still cover costs. For, owing to reduced output, his overhead cost rises, and therefore his average total cost. If he is to survive without abandonment of plant he must devise some new gadget to attract buyers. This may mean new equipment in the factory to produce the improvement, new plans for financing the purchase of cars, new advertising, more liberal discounts to sales agencies, all of which interfere with costs, usually raising them.

It is evident that when instability of demand (whether produced by the seller or resulting from conditions beyond his control) is prominent in any field and when the type of production employed necessitates the use of large amounts of fixed capital, correspondence between price and cost of production becomes extremely uncertain. The description invented many years ago to characterize the production of steel, "a feast or a famine," comes to be more or less applicable to many sorts of production. Producers strive to prevent this uncertainty. They attempt to forecast the demand, and they employ experts to calculate the costs of various amounts of output. Having made these forecasts, the producer is likely to set a price on his commodity and then sell what he can,

striving always to make the sales equal the output forecast. Such prices have come to be called "administered prices." Doubtless, too, the managements of the large producers of motor cars, cigarettes, popular brands of men's clothing, electric refrigerators, and many similar goods hesitate to inaugurate price wars. Some of these commodities sell for long periods of time at stable prices, which fact gives rise to the impression that the producers have a considerable degree of monopoly power.

The existence of monopoly is constantly denied by representatives of the management of these concerns. That a monopoly such as that possessed by a public utility exists in the production of these commodities is unbelievable. There is rivalry among the different producers. The large amount of advertising carried by the producing concerns in magazines of large circulation shows, by contrast with the very infrequent advertisements of public-utility concerns, that a striving for the maintenance of a share of consumer purchases is going on in the one case and not in the other. Yet such administered prices are certainly not the type of free competitive price that prevailed in the world wheat markets before government interference became common.

QUESTIONS AND PROBLEMS FOR DISCUSSION

1. Which prices should you expect to rise most rapidly during a boom—those charged by public utilities, those prevailing in purely competitive markets, or those charged under the system of administered prices? Which should you expect to fall most rapidly during a depression?

2. Monopolies generally have been regarded as antisocial. Why?

3. Name several commodities the prices for which are set according to the "follow the leader" principle. By what means has the leader obtained supremacy?

4. If you had your choice and were interested solely in making money, should you prefer a monopoly of (1) a good for which the demand curve was very steep or nearly flat; (2) one for which the demand was stable or unstable; (3) a luxury or a necessity?

5. It is often argued that because manufacturers have adopted administered prices in so many instances, production by farmers also must be limited. What sort of price system should you expect if all producers limited output at will? Would production increase or decrease? Would the income of the people of the United States be increased or decreased?

CHAPTER XIII · The Dynamics of Competitive Price



THE POSITION OF THE INDIVIDUAL PRODUCER UNDER STATIC COMPETITION

With this chapter we pass to the consideration of certain problems of price under competition and changing conditions. This and the succeeding chapters are concerned with what has come to be called "dynamics." In all the preceding chapters of Book II, except where specifically noted, it has been assumed that no great changes in the methods of production were taking place. It has been assumed also that demand for commodities was stable unless it was manipulated by sellers. Before the effects of changes in methods of production and in demand are discussed, a summary statement of the position of the individual producer under relatively changeless or static conditions will serve to bring out the contrast with the effects of such changes.

If competition were free, if there were no restraints on the entrance of firms into and withdrawal from any line of production, if the agents of production could move freely and without friction, and if there were no changes in demand or in methods of production, then normal price would always prevail. The individual producer, under these conditions, would operate in a world where economic production could be carried on with something like the same precision as routine tests in a chemical laboratory. Mistakes might happen as the result of negligence, incapacity, or willful error on the part of individuals, but not because of uncertainty in the economic environment. Technique, costs of the agents, and demand being known, all production would become mere routine. Under such conditions there would be no entrepreneurs, but merely directors or supervisors. If the markets were large enough to permit many firms to sell in each of them, prices would be competitive; if they were not, partial monopolies might exist. If markets are competitive, then costs will be the same for all producers. This must be so, because all have equal access to raw materials, to skilled, unskilled, and directive labor, and to the market for capital. If, however, anyone has special access to raw materials, such as the ownership of a superior

mine or farm or site for retailing, that mine, farm, or site will be valued according to its superiority, and if the favored producer keeps his cost accounts by the rule of opportunity cost, he will find that he must charge as part of his cost of production a rent for the superior productiveness of such resources. So too with managerial labor and skilled labor. Exceptional mechanics receive wages commensurate with their productiveness, and the managers, whether hired or self-employed, also are valued in proportion to their contribution to the product. In this imaginary world all agents of production cost what they are worth in the productive process.

The individual firms may not be of the same size although, if the terms "equal access to raw materials" and "equal access to markets" were strictly interpreted, all might be exactly equal in size in any particular industry. It is perhaps pushing the assumption of equality too far to suppose that all would produce exactly the same amount. Yet it has been found that in a certain area where topography and soil are uniform farms are mostly of the same acreage, though farming is by no means carried on under the conditions assumed above. At any rate it is necessary to insist that lowest total unit cost is the same for all.

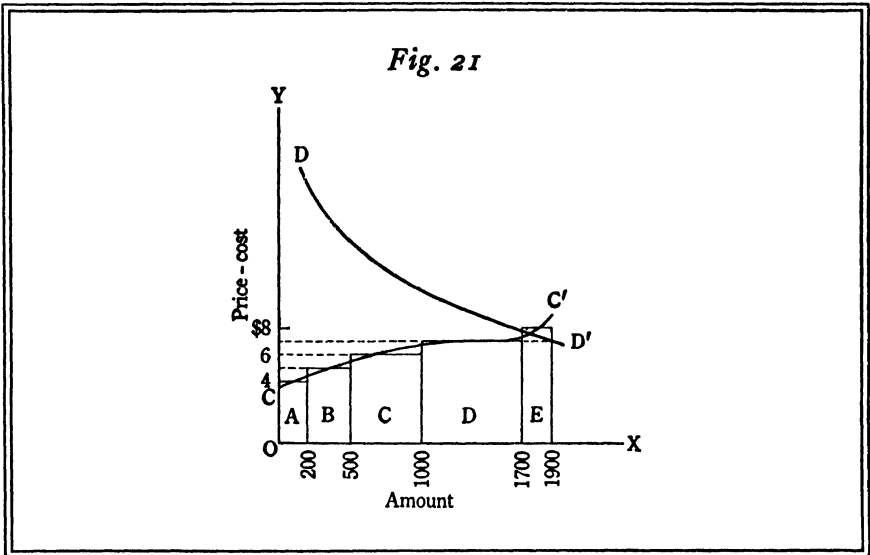
Of course it is highly unreal to assume that all producers have equal access to the markets and equal access to raw materials and directive ability. It is unreal also to assume that the demand for products can be forecast with a high degree of certainty. The agents of production are not perfectly mobile, as has been noted frequently in the preceding chapters. Because methods of production change rapidly in economic life, as we know it, machines, railroads, and buildings become obsolete. Some of the effects of changing conditions were explained in Chapter X in the section headed *Prime and Supplementary Costs*, but many have not been considered.

ENTREPRENEURS' DIFFERENTIAL COSTS

It is a matter of common knowledge that even in relatively stable times entrepreneurs make business profits ranging from large losses (minus profits) to very large positive amounts. In part the latter are caused by monopolistic conditions. But they also persist in businesses where little of what is commonly called monopolization exists. Not only the existence of imperfect competition, but also the forces of progress and decay give rise to differences in the profitableness of investments made by different concerns in the same line of industry. Another factor

that affects costs and therefore profits is the varying ability of management. The ways in which management can succeed or fail in the struggle to keep costs down and prices up are very numerous. In a stable, competitive world all estimating of future demand, future costs, and the progress of improvements would be absent. But in the real world they are not.

In Fig. 21 the relative positions of five groups of entrepreneurs have been illustrated. The most efficient firms can sell at a price of \$4 and



still cover their costs of production ; Group B can sell at \$5 ; and the least efficient group, E, can cover costs only if they sell at \$8. Since it is evident that price cannot agree with the costs of all the different groups, it is necessary to explain the relation between cost and price under the assumption that these differences in cost exist. What that relationship will be depends on the different conditions that may prevail within the industry where the differentials are found.

The temporary price at any time will depend, of course, on the conditions of demand and supply. In the figure it is \$7. But if the demand curve moves to the left and downward, price may fall to \$6, the cost of production of the moderately low-cost firms. There will be overcapacity and overproduction. Gradually the sellers whose costs are above \$6 will be forced to stop producing because they are unable to cover their costs of production. But, as the figure illustrates, they will

probably not do so at once. Seller E is not covering all costs of production, but he will remain for a time, provided he can cover all variable (prime) costs and earn some additional surplus to compensate in part for his own time and effort and for his investment in fixed equipment. Ultimately, if there is no further change, the exodus of the less efficient firms will restore a better correspondence between cost and price for most of the firms.

If demand increases suddenly, all firms will probably increase output, but for a time price will rise much above the costs of even Group E. Then new firms enter the field, supply is increased, and price falls again to about the level of the costs of Group D. It may be asked why additional firms having costs above those of Group E do not make up the new supply. The answer is that when, after the rise in demand, the profits become great, labor and capital come into the industry. They are brought in by entrepreneurs, and these entrepreneurs are most probably distributed among the grades of efficiency represented by the line CC' .

We can now answer the question raised in the preceding paragraph: What will be the price at any given time? The answer is the scarcely satisfactory one, that supply and demand will determine the price. But the supply, under changing conditions, is not closely connected with the cost of production of any firm as it would be under perfect competition and static conditions. Some types of business, which are easily entered by new men or which do not require great amounts of capital, seem always to be in a chronic state of overexpansion and overproduction; that is, a large part of the supply is from firms in Group E. The retailing of groceries has been one of these. Investigations show that in all cities the turnover of individual ownership is always very great. Anyone who has observed conditions in his own neighborhood has had ample proof of this condition. It is not easy to explain the constant supply of small investors who thus hazard their savings and their own earning power in such a precarious undertaking. Among the factors contributing to the situation are the desire of the hired employee to become his own boss, natural ambition coupled with less well-developed power of self-appraisal, the competition of wholesale houses to create new outlets for their goods, and lack of knowledge of the probable competition facing a new enterprise.

Price, therefore, in most competitive businesses, tends to be fixed by demand and supply. But the supply is brought to market by concerns that have different average costs of production. That is, the volume

that they are able to sell is such, and the price is such, that some make large profits, some merely cover costs, and some, at any given time, incur losses. The price at any given time is not likely, therefore, to agree with the costs of either the lowest-cost or the highest-cost sellers. It agrees more nearly with the costs of the large middle group. In Fig. 21 this is Group D. The tendency of price to agree more nearly with the average costs of the large middle group rather than with the costs of the most efficient or the least efficient has been found by actual investigation of industries many times. In fact in no case where conditions have been substantially competitive has any other result been found.

The explanation of the agreement of price with the costs of the large middle group rests on the phenomena of risk and uncertainty and on the equally important phenomenon of the distribution of ability among men. Risk and uncertainty are involved in the following way: If a large number of people are asked to *estimate* (not measure) the effect of any change upon some business with which they are familiar or to estimate the distance between two buildings, it will be found that their guesses will differ. Some will guess with a high degree of accuracy, and some will make guesses that are far from the actual facts. Between these two groups another and much larger group will estimate less accurately than the best but not so inaccurately as the worst. All phenomena in which chance and estimate are involved show the same distribution. If many persons pitch pennies at a small circle on the floor a few will register direct hits, and some will be wide of the mark, but the maximum density will be found to lie in a zone more or less close to the circle.

One function of the entrepreneur in a disturbed and changing world is to estimate the value of many factors and the occurrence of many events that cannot be forecast accurately. For example, a trader in wheat may know with a considerable degree of accuracy what have been the rainfall and the temperature readings over the entire wheat-growing area of the United States. But this knowledge will not enable him to forecast the wheat crop accurately. Many other factors affecting yields cannot be measured with the same degree of accuracy. Examples are the prevalence of pests and the percentage of germination of the seeds. It follows that any estimates of the wheat crop upon which the traders base their bids and offers for future delivery are distributed in somewhat the same pattern as the pennies that were pitched at the circle on the floor. In such enterprises as retailing, wholesaling, manufacturing, mining, and transportation many estimates must be made. A retailer

must estimate the probable volume of his trade when renting a store building; he must guess at the tastes of his customers, the movement of wholesale prices (to enable him to determine when to lay in large inventories and when to buy "from hand to mouth"); and the reliability of those to whom he grants credit.

The supply of any commodity requires some preliminary acts on the part of the producer. Farmers must plant wheat from five to ten months before harvest, manufacturers must buy materials, employ labor, contract for supplies, and take many other steps in anticipation of the actual delivery of the good to the buyer. The amount and kind of commodities they produce depend on their estimates of future conditions. The amount they produce and offer for sale, together with the demand, determines the price at which the commodities sell. If all estimates were as accurate as the finest measurements made in laboratories, the entrepreneurs' differential cost curve would show less difference between most and least efficient, and price would agree more nearly with the average cost of each producer. But the estimates cannot be made with any such degree of accuracy. The producers whose estimates happen to be best will make profits; those whose estimates happen to be worst will suffer losses. But the main part of the supply will come from some more typical group—such as the D group in Fig. 21. Hence the price is more likely to agree with their costs.

Another factor besides chance affects the costs of the entrepreneurs. It has been known for many years that if any large group of people is set to perform some task where performance can be measured accurately a distribution of achievements similar to Fig. 21 will be found. Some will do the task very well; some will do it very badly; but the large majority will cluster around some central point between the two extremes. Individual differences are, therefore, the second ground of explanation of the entrepreneurs' differential-cost curve, CC' . In part, the ability of the entrepreneurs is the ability to estimate unmeasurable factors. Hence the risk factor and the ability factor are related. Another function of the entrepreneur is that of co-ordination and management. Here again individual differences play a part. The ablest will keep costs low because they have the capacity to co-ordinate very well, and some will have high costs because they do not possess even average ability, while the majority will cluster around some intermediate point. Now the costs affect the supply of the commodity and therefore the price.

It must not be assumed that all the producers in the A and B groups at any time remain there permanently. Owing to the elements of risk

and uncertainty, even the ablest entrepreneurs will make mistakes and thus find themselves at times in the C group and even on occasion in the E group. Investigations have shown that there is considerable shifting from year to year. There may be also a more or less permanent exchange of places between the members of the groups. Some of the producers in the E group may be new at the business, and when they have had more experience, and if they can survive, they will move into the lower-cost classifications. Likewise, some of the firms in the A group may lose their efficiency, owing perhaps to a change in personnel, and progress toward the higher-cost classifications.

COST DIFFERENTIALS AND IMPERFECT COMPETITION

In the preceding section it was assumed that the efficient producers in Group A possessed no advantages not open to others, except the single one of superior ability. Often, however, lower costs may arise from the inability of all producers to obtain access to all the advantages open to the low-cost producers. Some of the more important advantages are: (1) superior location, (2) access to superior raw materials, (3) locally lower labor costs, (4) secret processes, (5) patents, and (6) the patronage of an established clientele. That any or all of these advantages, if they are not accessible to a majority of the producers, will contribute to lower costs is obvious.

In computing the costs of those who have these advantages care must be exercised to make the proper distinction between opportunity costs and costs of record. In the following illustration, the proper distinction is made. Take the case of a farmer who has inherited a farm in a fertile region. The father of the present owner paid \$10,000 for the land and buildings forty years ago. This is the cost of record. The farm is still carried on the books of the present owner at \$10,000, though it could be sold for \$25,000. The latter figure is the opportunity cost. Because the owner computes his costs by entering among them an interest charge of 5 per cent on \$10,000, he finds that the produce of the land costs \$500 for the land use alone. Of course, other costs for labor, taxes, seed, animals, and depreciation on machines also will be included. But since the land could be sold for \$25,000, the costs are understated by the difference between 5 per cent of \$10,000 and 5 per cent of \$25,000, or \$750. As has been shown, the cost that influences supply, and therefore price, is not cost of record but opportunity cost. In this case, the degree of intensive use of the land, the extent to which the other factors of pro-

duction will be combined with the land, should be determined by its opportunity cost. Hence the supply of the produce of the farm will be determined in this case and in all similar cases by its opportunity cost, not by its cost of record.

Several other instances in which the same type of incorrect calculation might occur follow, and the reader can easily supply many more. An able businessman, who is both entrepreneur and owner of the productive property he manages, allows himself a salary of \$10,000 a year. But if he wished, he could obtain a salaried position as manager of some other business, which would pay him \$15,000. Obviously the cost of record is \$10,000, but the opportunity cost is \$15,000. A firm purchases a plant in good condition at a receiver's sale. It gets a bargain. The purchase price is entered in the capital ledger at the bargain figure and thus becomes the basis for calculating the overhead cost. But the plant could not have been constructed for so low a figure; nor would it have been sold at that figure by a firm that was conducting a successful business. Hence the cost of record is lower than the opportunity cost. If the correct opportunity costs had been used, the average cost of production in both cases would have been higher.

A further distinction must be carefully made between opportunity cost and capitalization of advantages arising out of imperfect competition. We may take again the case of the owner-operator who, on his books, arbitrarily sets his own salary at \$10,000 a year. Because he might have obtained \$15,000 a year as a salaried manager, it has been argued that he should charge the latter amount to cost of production. Let the assumption now be changed in one particular. Assume that he could get \$15,000 salary as a hired manager, and that his profits over and above all costs, both prime and supplementary (his own salary excluded), amount to \$20,000. It would not be correct reasoning for the owner to count as his salary, and therefore as cost of production, the entire \$20,000. In another case, a patent may enable one concern to produce much more cheaply than all rivals. It is valid for the firm to refuse to sell the patent for less than the capitalized value of the surplus earnings to which it gives rise, but it is not valid for it to compute costs so as to include the earnings of the special advantage derived from the patent.

Any special advantage arising from the imperfections of competition is valuable. But the earnings arising from such advantage are not to be counted as costs. Only competitive values that give rise to opportunity costs are costs in the sense that they are obstacles to be overcome in production of goods.

IMPROVEMENTS AND DIFFERENTIAL COSTS

In Fig. 21 the A group have lower costs than any other group. Sometimes this advantage is due to the use of improved methods that are not protected by patents or guarded by secrecy. Any concern may be free to adopt these new methods if it wishes to do so. But the A and B groups have put them into use, and the other groups as yet have not done so. The explanation of this inequality is not difficult. Possibly the two groups, A and B, are managed by more alert entrepreneurs, who are responsible for the superiority in this and other respects. Superiority in respect to estimating unmeasurable conditions affecting production cannot be taken away from these groups by competition, but their superiority in respect to methods can be.

When the improvements are first introduced, the superior entrepreneurs experience lower costs. If they also expand their outputs the supply will be increased, and more firms will be pushed into the group that is making losses. If they do not expand, then the supply will not be affected at once, and the lot of the inferior producers will not be disturbed. If the groups C, D, and E adopt the improvements quickly, the curve CC' will fall toward the OX axis, and even the firms in Group E may make a profit. But this profit will attract new entrepreneurs, who will probably be distributed along the entire curve. There will be more superior, more typical, and more inferior firms. Supply will be increased, and price will fall below \$7 a unit, but there will still be the same relative distribution of firms along CC' . The consumer gains the entire advantage of the reduction of cost that has been effected by the introduction of the improvement by the typical firms in Group D.

If the improvement is not adopted by the typical and inferior firms, new producers will come into the industry, and they will be distributed in ability along CC' . But they will all introduce the improvement, and the old firms that failed to make the change will be pushed beyond E and ultimately excluded from production through the merciless, yet socially beneficial, operation of competition.

When the efficient producers find that they can make larger profits by expansion after the improvement has been introduced, a condition of partial monopoly may result. It may happen that the new methods require very large plants in order to make full use of fixed equipment. If this is the case, the smaller firms and those not led by men of high managerial ability will not be able to make the changes required by the new methods. The A and B groups will expand and drive all others

out of the field. Then, with the reduction of the number of firms and the increase in fixed capital, will arise the problem of maintaining the markets of the few firms. Advertising and other methods of sales promotion will be made use of, and output will be controlled to maintain price. These are the familiar signs of the development of imperfect competition and administered prices. Price, as was shown in the preceding chapter, will not usually be as high as though a complete monopoly prevailed, but it will frequently be higher than the lowest total unit cost of the firms. Output will be less than the optimum. Price will not be an equilibrium price, and the factors of production will not earn the same return as they earn in the more competitive branches of production.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It is sometimes said that if the factors were perfectly mobile, adjustments to changes in methods of production would not result in profits and losses. Why?

2. Suppose that the five groups of concerns represented in Fig. 21 should organize a combination to control price. Suppose also that they were unable to eliminate any of the producers, what price would they probably set on the product?

3. Suppose that the combination in the preceding question set the most profitable price but could not prevent the entrance of new firms. What would eventually happen to the supply? to the price?

4. If each group among the various groups of firms represented in Fig. 21 can protect its market from the competition of other groups, will there be one price for the product? Will there be lower prices in the markets served by groups A and B?

5. Some of the codes introduced under the National Recovery Administration forbade the introduction of improved machinery. This prohibition was admittedly temporary. Would the prohibition against better machinery remove the differentials? Would it prevent the existence of the "excess capacity" represented by Group E?

6. Suppose that a beneficent despot should by edict fix the price in Fig. 21 at \$8, the average cost of Group E. Assume that the demand remained as shown by DD' and fundamental conditions outside this industry remained constant. Could excess capacity be permanently prevented?

CHAPTER XIV · The Dynamics of Price: Varying Cost

THE EFFECT OF CHANGES IN DEMAND ON COST

In the preceding chapter the effects of certain types of change on the conditions of supply and on price have been considered. Most of these changes affected the different producers unequally. In this chapter we shall consider changes in demand that affect all producers more or less equally. It follows that here we are making different assumptions from those used in the preceding chapter: (1) It is assumed here that the producers all produce at substantially the same lowest total unit cost. (2) It is assumed that there are many sellers and many buyers. (3) It is assumed that competition is perfect, that it is free from governmental interference and from lags arising from unequal knowledge, custom, inertia, and the slowness of movement of the factors arising out of large amounts of fixed capital. Another way of saying that these lags are absent is to assume that enough time elapses to make the impediments to perfect competition here enumerated of no effect on prices. (4) It is assumed also that there are no improvements in methods of production. Later some of these assumptions will be withdrawn, and the effects of the changed condition will be considered.

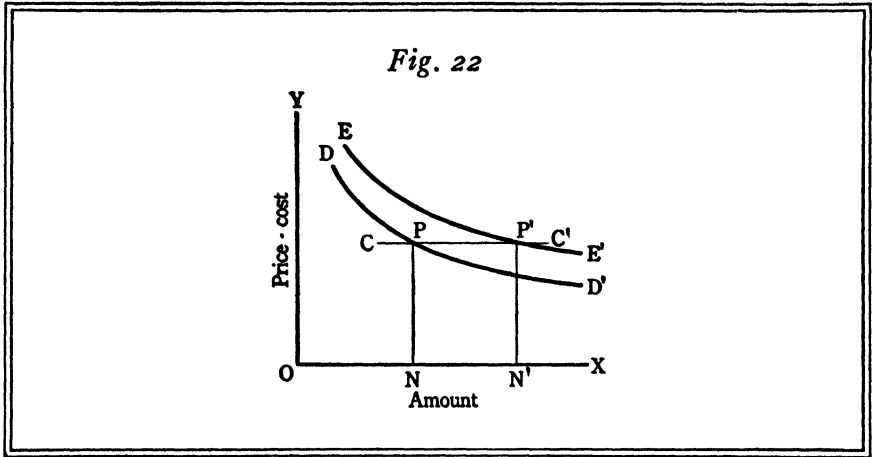
We now ask, What effect will an increase of demand that offers additional sales to all the competitors have upon price? It must be emphasized that the problem is considered here with the four assumptions enumerated in the preceding paragraph in force. In the following analysis three types of effect on cost are explained.

CONSTANT COST

When an increase in the demand for a commodity is met, after a time, by an increase in supply brought to market under exactly the same conditions as those prevailing before the increase took place, the commodity is produced under constant cost. In Fig. 22 cost for varying amounts of output is represented by the line CC' . Let it be assumed that with the demand curve DD' the numerous competing sellers are able to supply ON

units at a cost of P . This is a stable equilibrium price. Of course, the actual market price is sometimes above and sometimes below NP because production is never entirely stable.

Now let it be assumed that the demand changes so that the demand curve DD' comes to have the position represented by EE' . When the demand first begins to change, undoubtedly the price will rise above NP unless the change has been fully anticipated and prepared for by the sellers. But after adjustments have been made the price settles at $N'P'$,



which is equal to NP . The increase in demand has had no effect on price but only on the volume of sales, which has increased from ON to ON' . If we assume a decrease in demand, the price remains the same, but the volume of sales declines. Here it is evident that price is wholly controlled by cost and that demand determines only the volume of sales.

CONDITIONS NECESSARY FOR CONSTANT COST

A moment's reflection should cause one to doubt whether there is any good whose cost could remain the same if the demand should increase indefinitely. Eventually, we should suppose, raw materials would become scarcer, and cost would rise. Even if the raw materials were derived from cultivation and not from mines we should suppose that the principle of diminishing returns sooner or later would cause expenses of production to rise. On the other hand, if the demand should decrease greatly, we should suppose that the same principle would bring about a decline in costs. Finally, we know that either the increase or the de-

crease of demand may alter the conditions of specialization and division of labor within the industry. It will stimulate or retard improvements. Should this happen, costs would again be changed.

It is evident that there is no commodity the supply of which could be increased or decreased indefinitely at constant cost. Hence the line CC' cannot be extended to the OY axis or indefinitely to the right. We are led to the conclusion that constant cost is a condition that can exist only under certain special conditions and, even then, only for a limited variation of demand. In general these conditions are such that costs are not affected by either an increase or a decrease in the demand for the primary factors of production—land (raw materials), labor, and capital.

The principal conditions that give rise to constant cost are as follows: In the first place, the raw materials used in production must be so plentiful relative to the demand for them in this industry that a considerable change in the demand for the products of this industry will not cause the factors to become either more or less scarce. This happens when the supply of the material is freely reproducible, as are many textile fibers, and when the demand for the material for the manufacture of the good in question is very small in comparison with the total demand for it. In the second place, the labor necessary to fabricate the good must be forthcoming at constant wages. This again can usually be true only when the demand for labor in this industry is a very small part of the total demand. In the third place, the industry must be one in which the optimum size of plant can be reached easily, so that many producers can serve the same market. In the fourth place, the methods of production must be stable. That is, it must be impossible to improve efficiency by adapting methods and machines from other industries.

These requirements are so rigid that it appears improbable that many industries can operate at constant cost over wide ranges of output. Statistical verification of these assumptions is next to impossible, because we are concerned here with adjustments that take place under conditions of perfect competition and in the absence of changes in the arts, and these are conditions seldom encountered in real life.

Constant cost as it has just been described should be clearly distinguished from a level price trend for any commodity. The latter may be found statistically in periods of price stability, when the index numbers or general level of prices will be found to vary little over a number of years while the volume of production is steadily increasing. The phenomenon of constancy of price is found also in the case of small articles sold at customary prices. It may be asked whether constancy

of price does not show constancy of costs, with minor fluctuations arising from temporary deviations from the equilibrium or normal level of costs. Of course this might be the case, and it probably would be if we could assume perfect competition and no change in the arts.

Actually constancy of price arises from a variety of conditions. The producers may not be able to compute the costs of a single product because several products may have so large an element of common, or overhead, costs that the isolation of the expense of producing one of them is either impossible or not worth the expense of the necessary analysis. When this is true the producers turn out the group of wares, taking care that none sells for a very long time at a price that does not cover prime costs, but without knowledge of exactly how much the total cost of each is. They expect to make enough to cover all costs and perhaps, if they are fortunate, more than normal profit each year.

When an index of the general price level shows little variation it is never true that all prices are constant. There is always some movement. But these movements offset each other. The same may be true of entrepreneurs' costs. Wages may be rising while the costs of raw materials are going down. Inventions may cheapen the production of raw materials in spite of the operation of the principle of diminishing returns. Total costs and even variable costs remain at about the same level, but this constancy is due to numerous offsetting changes. Hence the trend of price shows neither rise nor fall.

Still another condition may give rise to a level price trend. It sometimes happens that the demand shifts from one commodity to another commodity that is composed of many or nearly all of the same materials and that uses the same types of labor and about the same amounts of capital as the good from which the demand has shifted. For example, the railroads were the great users of steel from about 1870 to 1920. Then, during the third decade of the century, much larger numbers of people began to travel by automobile and shippers began to make use of motor vans. The demand for passenger service and for some types of freight service declined. New railroads were no longer required. The automobile industry began to use the steel that formerly had been used by the railroads, and it doubtless attracted many workmen who otherwise would have been employed in building new or improved lines of railroad track or in constructing cars and engines. In agriculture the relative decline in the demand for heavy foods such as meat and cereals as the population became more urban and less rural, as machines took over many heavy tasks, and as new tastes and new

ideas about diet developed, was more or less counterbalanced by an increased demand for dairy products and fresh vegetables and fruits. The same land that formerly had grown wheat in Minnesota was turned to corn, fodder crops, and pasture, and some of the farms of Iowa and other states began to raise dairy herds instead of beef cattle.

When such shifts of demand are accompanied by shifts of the factors of production from the industry where demand is slackening or receding to the industry where demand is growing, it is possible that there may be no change in the prices of either the good for which the demand is declining or that for which it is increasing.

The fact that the output of an industry has increased while prices and costs have remained the same over a considerable period of time is no proof that the industry has been operating at constant cost. In only one of the illustrations given in the preceding paragraphs has "constant cost," as the economist understands it, been in operation. That is the case where the increase in the demand for one commodity represents a shift from another commodity made with the same materials, labor, etc. Even in this case a practical illustration cannot be found. A level price trend is not proof of the operation of constant cost. Usually where price history seems to indicate constant cost, dynamic conditions will be found present.

CONSTANT COSTS AND THE SHIFTING OF COSTS

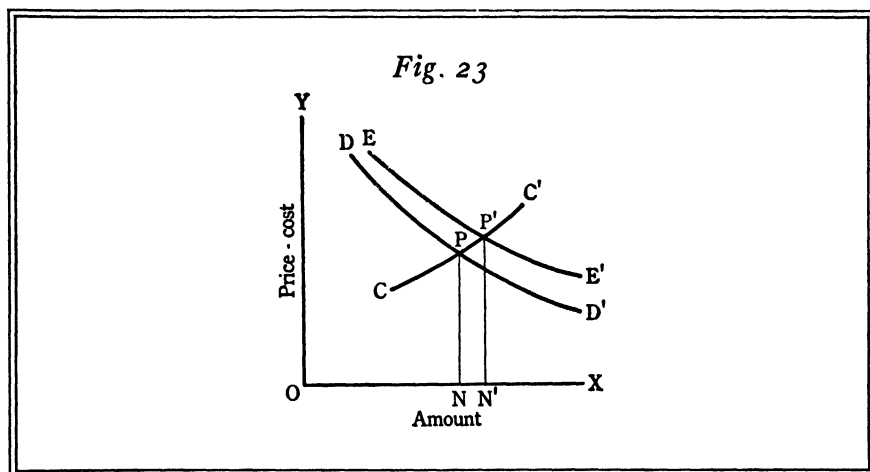
In the discussion of interdependent prices it was shown that a given increase in the cost of one of the factors of production may cause an equal rise in the price of the finished good to which the factor contributes, may cause the price of one or more of the other factors combined with the one factor to fall by an equal amount, or may cause the price of the product to rise somewhat and the prices of the other factors to fall somewhat. If constant cost prevails, an increase in the expense of production will affect the price of the finished product only. Since constant cost occurs only when the factors of production have many alternative uses and when, in consequence, the demand for them to produce the good in question is relatively very small, it must be supposed that when one of these factors rises in price the rise is common to all industries. Hence the price of the good must rise by the full amount of the increased cost if the good is to be produced at all. The rise in cost may be due also to a new, or special, cost attaching only to this industry.

If, however, costs only appear to be constant, and if the level price trend is due to offsetting changes or to control of prices by custom, a rise in the cost of one of the factors will have an indeterminate effect. The various reactions to such a rise are numerous. Sometimes the producers believe that the rise is temporary, and they fear to permit the price of the finished good to go up lest substitutes become established. Or they find that high profits in the production of the finished good have attracted so many competitors that there is an oversupply, and prices cannot be raised until some of the competing firms shall have been expelled. Sometimes a considerable rise in the cost of one of the factors is the signal for an almost concerted raising of prices by all producers. This takes place when producers have been squeezed between the slow advance of the costs of other factors and a relatively rigid and inelastic demand at and near the ruling price. The rise in the cost of the one factor determines the producers to raise the price. Often the opportunity to do this is given by a revival of prosperity or by the bankruptcy of some of the competitors. In general, however, any increase in the unit cost will cause prices to rise by the same amount.

INCREASING COST

A good is said to be produced at increasing cost when a rise in the demand for it is accompanied by an increase in total production, which increase cannot be put on the market except at a higher cost than prevailed before the increase took place. In Fig. 23 the operation of the principle is illustrated. Since the phenomenon is frequently encountered in agriculture, the illustration is given in terms of wheat. Again we shall assume that when ON units are produced the price NP is a stable equilibrium price; that is, the factors of production earn as much in the production of wheat as they could earn in any other use. If the demand increases from the position DD' to that of EE' , what will be the effect on the cost of production? Again, as when constant cost prevailed, price will probably rise temporarily much above NP , the old producers will expand their production, and costs will rise above NP . But in this case a part of the rise will be permanent. This will necessarily be true because the land used by the farmers who formerly supplied ON units will be cultivated more intensively and, owing to the operation of diminishing returns, both marginal and total costs must be higher than before. But although total costs rise, they will not rise as much as price, because the marginal-cost curve lies above the total-

cost curve when any business unit is expanded beyond the point of lowest total unit cost. Wheat-growers will make profits. These profits will attract additional producers, who will use, for the growing of wheat, land that had either been used for grazing or for the production of some less profitable crop. During the period 1914 to 1919 large areas of new land were brought under the plow and seeded to wheat both in Canada and in the "dust bowl" area in the United States. In many cases farmers broke up the native grasslands in older areas. These lands were often less suitable for cultivation—rough areas where erosion both by wind



and rainfall was rapid and low-lying areas that lacked adequate drainage. The new lands in Canada, however, as well as in Argentina, where also there was an increase in wheat-growing, were often more fertile or were more favorably situated in respect to climate than some of the older lands. In the United States, owing to the introduction of larger and improved machines and the use of motorized equipment, production on some of the new wheatlands was less costly than on some of the older lands. In other words, the actual conditions did not agree exactly with the assumptions necessary to bring about increasing cost. But on none of the newer wheat areas could production be carried on as efficiently as on the best of the older areas.

The contrast between the conditions giving rise to constant and to increasing cost must be clearly grasped. In the preceding sections of this chapter, where the effects of increased demand were discussed, the assumption was made that the increased use of the factors to produce more of any given good did not make them appreciably scarcer. The

new producers who came in when the demand increased were able therefore to bring the commodity to market for the same average cost as the old producers of the same personal efficiency. In the case of wheat the earlier assumption is not valid because the land factor is important in production and because wheat is an important part of agricultural production. Obviously, however, if the demand for wheat increased while that for beef decreased, and if the land left unutilized by the decline in the demand for beef were equally good for the production of wheat, it would be unnecessary to resort to poorer lands to meet the increased demand for wheat. Here it is assumed that the demand for wheat is not offset by a decline in the demand for other farm produce that could be grown on the same land and with the same intensity of cultivation.

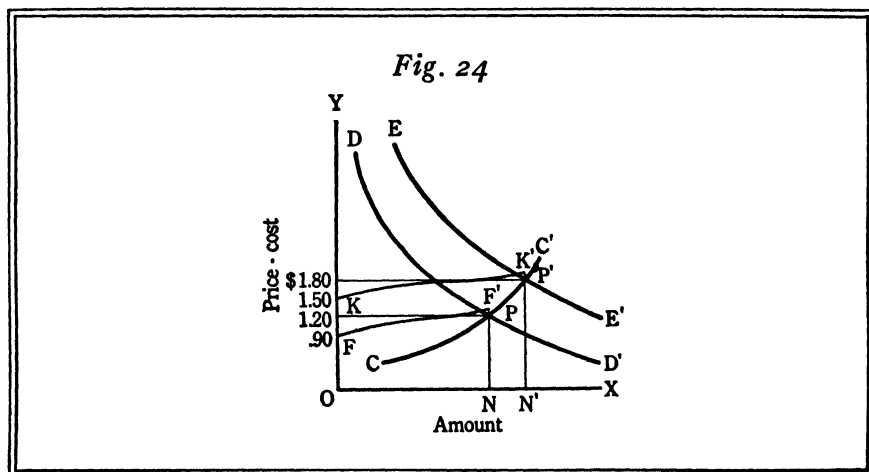
Because the new producers cannot grow wheat on the poorer lands as efficiently as the old producers on the better lands, the cost of production will be higher and the necessary supply price higher. Hence the equilibrium price must be higher. But this higher price, $N'P'$, will be paid to the old producers also. At first these producers will find that a profit results from the new and higher price. Each will tend to push production to the point where the marginal cost and the price agree. This will raise the average cost of production on even the better grades of land. But there will still remain a profit for the farmers who produce on the better lands. In time, however, competition among tenants will raise the rents of these lands, because the tenant who has occupied the poorer land will find he has made no profit, or a smaller profit, and will try to obtain the use of better land. He will compete with the old lessees of that land and force up its annual rental value.

When the better land is farmed by the owners there will be no explicit rent. But the farmer will find his profits have risen. If, however, a prospective purchaser appears, the farmer who was in possession of the land before the demand increased will hardly part with this property unless he is paid the present worth of the added profits that can be made from its use in the future. The value of the land will rise. When a new owner who intends to farm the land takes possession, he will find that his costs, including the annual charge for interest on investment in land, have risen. They will be approximately the same as the cost of growing wheat on the poorest land in use, the differential costs arising from differences in the ability of management excepted. The cause of the rise in the price of the land has been the increased rental, or productivity, value of the land. The same process will cause the price of land let to tenants to rise.

The line CC' represents the highest, or marginal, costs of producing varying amounts of product when rent has been excluded. It shows that when ON units are produced, the cost of production on the poorest land, for which it is assumed no rent is paid, is NP . As the demand moves from the position DD' to the position EE' , poorer and poorer land is taken into cultivation, and old land is cultivated more intensively, which means that more capital and labor are used with the same area of land. The cost of production of the additional supply required to satisfy the growing demand becomes higher and higher. The movement of this marginal, or additional, cost as the entire supply is increased is shown by CC' . In terms of geometry, CC' is the locus of the points of marginal cost for amounts of product varying from zero to ON and ON' .

ENTREPRENEURS' DIFFERENTIAL COST AND INCREASING COST

The farmers who lease land from the owners cannot very long retain the profits arising from an increase in the demand for agricultural products. The behavior of entrepreneurs' cost when increasing cost prevails requires further explanation. In Fig. 24 the effect of an increase



in demand upon entrepreneurs' cost has been illustrated. If competition is perfect, and if all producers have the same ability, the costs of all will be (at the point of lowest total unit cost) equal to \$1.20 a bushel, rent included, when ON units are produced. If, however, competition is not perfect, and if allowance is made for differences in ability, costs will vary according to the line FF' . The costs will vary according to slight

advantages and according to ability. When demand moves from DD' to the position EE' , entrepreneurs' costs will be represented by KK' , or by a straight line parallel to OX , extending from OY to the intersection of CC' with EE' at P' , if perfect competition and equal ability are assumed.

It must be clearly understood, of course, that when the output is at ON' , the typical producers as farmers incur costs of \$1.80 a bushel and that not even the best farmers can produce at the old costs. The landlords (or the farmers in the capacity of landlords, if they own the farms they operate) will receive higher rents.

FORCES COUNTERACTING INCREASING COST

Thus far we have considered the effect of increased demand upon the cost of obtaining the products of farm, mine, and forest, under the assumptions that there were no discoveries of new resources and no inventions. These assumptions may now be removed and conditions brought nearer to reality. The general effect of any of these changes is to reduce the labor cost and the capital cost of obtaining raw materials. Wheat that formerly cost \$1.20 a bushel may be grown for 80 cents after an improvement in farming occurs. Moreover, even though the reduction in cost occurs at the same time that the demand for the product increases, it is entirely possible that a larger crop may be grown at a lower expense of production. Similarly, if the growing population discovers and cultivates new and better lands (as was the case with the surplus population of Europe during the eighteenth and nineteenth centuries) instead of cultivating the same areas more intensively, the cost of the product may actually decrease. Finally, if science reveals ways and means of tapping resources formerly unavailable, the additional population may be able to supply itself with the products of the field, the forest, and the mine without any rise of the existing normal cost.

If we were able to obtain comparable data of the cost of growing wheat or any other crop in any country over many years, we should probably find that during some of these years the cost of production was steadily rising, in others practically unchanging, and in still others steadily falling. But in every period of either falling or unchanging costs we should find that some of the conditions assumed in the statement of the law of diminishing returns had been violated and that, as soon as the assumed conditions were again restored, the cost of obtaining additional supplies had again begun to rise.

If the demand for raw materials and foods rises continuously, increasing cost will necessarily make its appearance unless the counteracting forces of invention and discovery are continuously accelerated. For example, we may assume that the cost of producing the marginal increments of the supply of wheat is \$1.20 a bushel and that annually the growing population requires 2 per cent more than in the preceding year. If there are no improvements of any sort, the cost of getting these annual additions to supply will certainly rise. But suppose that scientists discover a new variety of wheat that either yields much more than old varieties or can be grown in areas where formerly lack of rainfall or the shortness of the growing season prevented the cultivation of wheat. The pressure of intensive cultivation on the old wheatlands may be relieved, some of the poorest may be abandoned, the marginal cost of growing wheat may fall, and the price of wheat may fall. If we assume that with every increase in the demand for wheat an invention or a scientific discovery appears and increases the productivity of existing wheatlands in the same proportion, the principle of diminishing returns may be entirely counteracted, and the new demand may be met by an additional supply, which the producers will be able to put on the market at a constant marginal cost.

The prices of the products of agriculture to the final consumer consist of the expenses of growing them on the farm, of processing them in factories, and of marketing. The first of these is sometimes less than half the price the consumer pays for the finished good. A reduction of the costs of processing or of marketing reduces the price of the finished good, and it therefore is entirely possible that an increase in the demand for certain products may be accompanied by increasing costs of production of the raw materials on the farm and at the same time by reduced costs for transportation, merchandising, or processing. If these reductions are great enough, the consumers are able to secure larger amounts of the finished good at lower prices in spite of the fact that the cost of producing the raw materials has increased. For example, bread may fall in price while the cost of growing wheat increases. When the principle of increasing cost is used in the analysis of concrete problems, it is always necessary to distinguish the different stages in the complete productive process. Diminishing returns may bring about increasing costs in one stage, and inventions may reduce costs in succeeding stages.

But the forces that offset the effects of diminishing returns in the production of food and raw materials can hardly be expected to keep pace exactly with the rising demand for these products. A rise in demand

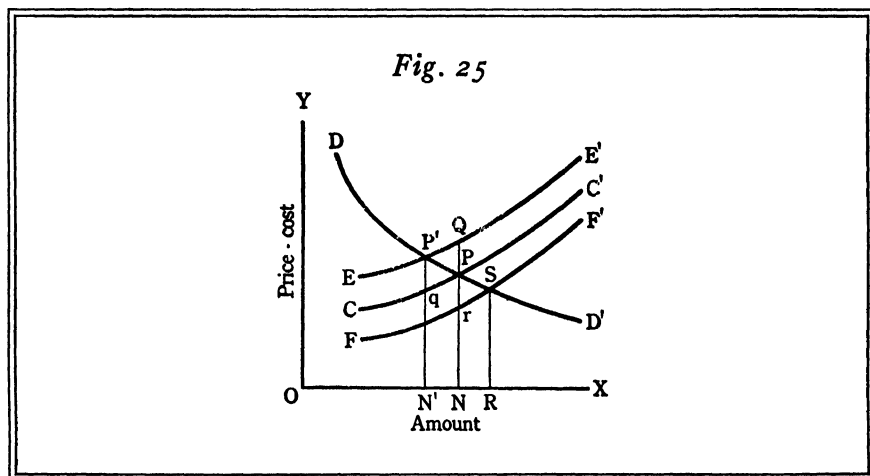
usually causes the price of one or more of those products to rise. If this rise in price stimulates improvements and discovery, if new resources are found, or if improved methods are devised, either for agriculture or for processing and marketing, price retreats toward its original level. Very often improvements outrun the rising demand, and price actually falls below the point at which it stood before demand increased. When these improvements occur in agriculture itself, the result is likely to be overproduction of farm products, and therefore many agricultural producers may find that they cannot earn the incomes they previously had earned; there follows a period of distress and discontent in the rural districts of the areas affected by the improvements.

Historically, the picture we have drawn is characteristic of many sorts of agricultural production, iron-mining, and coal-mining. Is there, then, any justification for saying that the extractive industries operate under conditions of diminishing returns? Even though costs are declining, may not the law of diminishing returns be in abeyance? The answer is that the tendency is always in operation. In the first place, we must note that were it not for the operation of the principle of diminishing returns, the cost under the improved conditions would have fallen even more than it did. In the second place, we must insist that the principle of diminishing returns and its consequence, increasing cost, are of the nature of universal physical laws, whose effects for a time may be obscured or counterbalanced by the ingenuity of man, but whose operation can no more be abolished or wiped out than can the law of gravitation. If the demand for wheat increases, its cost will begin to rise the moment man relaxes his efforts to perfect the arts and to delve deeper into the truths of science.

INCREASING COST AND THE INCIDENCE OF NEW UNIT COSTS

Under conditions of free competition, any expense necessarily incurred by the typical producers must be covered by the selling price of the goods they produce. Any increase in expense will add just that amount to normal cost. In the preceding chapter we saw that any increase in the unit expense of bringing to market a commodity produced under conditions of constant cost would raise by the same amount the normal price of that commodity. When, however, a good is produced at increasing cost, a given addition to the expense will raise the price by something less than the addition. The reason for this is evident from Fig. 25. In this figure CC' represents marginal cost of production when supply is varied along OX . When ON units are produced, the cost and the price

are NP . If one of the variable factors should rise in cost so that for any given amount of output the unit cost was greater than before by PQ , the new total cost of varying amounts of the product would be represented by EE' . Producers, of course, would try to shift the new expense to the consumers. With the demand represented by DD' , they could do so only by reducing output from ON to ON' . But when output has been so restricted, it is evident that the new expense is added to a smaller old



expense than when ON units were produced, since $N'q$ is less than NP . It is evident likewise that the new price $N'P'$ exceeds NP by an amount less than PQ .

The validity of this diagrammatic demonstration for real life is apparent when we recall that the producers, in their efforts to shift the new cost to the consumers, must reduce output. But reduction of output means the abandonment of the poorer sources of supply and the application of less of the variable factors to the fixed factor—land and other natural resources. Stated in other terms, the efforts of the entrepreneurs to shift the new expense result in the establishment of a new and lower margin of cost. The foregoing conclusion holds for taxes on the output of farms, forests, and mines. It holds also for additional labor costs and capital costs that bring about variations in the unit cost of goods.

It is sometimes supposed that any sort of subsidy paid to producers cheapens the product by the amount of the subsidy. Now a subsidy amounts to a reduction of the expense that the producers must charge the consumers in order to remain in business. Suppose that the subsidized product is being put on the market at a marginal cost of NP , in terms of

Fig. 25. The government now agrees to pay the producers rP for every unit sold. The effect of this reduction in expense (rP) when increasing cost rules in the industry is shown by the curve FF' . Since the reduction of cost gives the producers greater profits, they will expand the output until the marginal cost again is equal to the selling price. They will increase supply from ON to OR , and price will be RS . By the same reasoning as in the case of the added expense, we conclude that the price will fall by something less than rP . The economic argument also runs in the same terms. The producers, in order to expand the supply, must make use of poorer natural resources, and they must cultivate the old resources more intensively. Both procedures will lead to higher costs, and the subsidy therefore will be subtracted from a new and higher expense than that existing before the subsidy was added.

In the application of the foregoing reasoning to actual cases, certain cautions must be observed. What has been said holds only if the increase in expense or the subsidy is borne or paid in accordance with some recognized unit of output. If the cost of one kind of farm machinery should rise, we might discover some, probably a very small, effect on the price of the crop in the production of which the machinery was used. But it is not necessarily true that the enhanced price would follow the principles explained in this section. Farmers might discard the machine entirely and put another in its place; they might be able to economize in its use; they might change slightly the type of product grown on the land. A further caution is necessary when we apply this reasoning to any argument concerning the effect of subsidizing agricultural products. It has been said that such an aid would cause poorer land to be employed with a resulting increase in total product. This, in turn, would lower prices and therefore leave the farmers no better off than before the aid was granted. This conclusion may be in error in two respects. In the first place, the expansion of the farming area would be checked by increasing costs before price descended by the full amount of the subsidy, so that some benefits would still remain with the producers of all the product except the marginal increments. In the second place, the conclusion that the farmers would expand production depends upon the unstated assumption that they are already covering their expenses. If they were not, they probably would not increase their output. At any rate, they would not do so until the subsidy was made greater than the difference between expense and selling price.

THE MEANING OF "DECREASING COST"

In general terms, "decreasing cost" means that when the demand for a given product increases, forces are called into action that cause the cost of production of that article to be reduced. It should be noted that decreasing cost is in some respects the same type of phenomenon as constant or increasing cost. It is an effect on cost that follows from an increase in demand, just as are constant and increasing cost, but there are significant differences. The essential cause of increasing cost is the greater scarcity of one or more factors, usually natural resources. When costs remain constant under the pressure of rising demand, it is usually because no factor of production is made substantially scarcer by that increase. Now, while it is conceivable that an increase in the demand for a certain product would not immediately cause its constituent factors to become scarcer, it is hardly conceivable that an increase in demand would cause its constituent factors to become more plentiful. Increased consumption and use, if carried far enough, cause greater scarcity; they do not cause greater plentifulness.

It is clear, then, that decreasing cost cannot be the opposite of increasing cost when we are considering the production of an entire industry. Another way of stating the same fact is to say that the curves of variable cost and total cost that apply to a single plant or to the operations of a single entrepreneur are not applicable to an entire industry. Fig. 2, Chapter V, which is basic for all the value explanations, is explanatory of the conditions confronting a single entrepreneur. One of the assumptions underlying that figure is that the prices of all the factors of production remain constant. The entrepreneur pays no more for a day's labor or a ton of raw material when he produces a large amount than when he produces a small amount. Total costs in the entrepreneur's plant fall until the point of lowest total unit cost is arrived at, and then rise. This happens because the fixed plant is very inadequately utilized when only a small amount is produced and overhead costs are high. When production expands, the overhead cost remains constant and becomes a smaller unit charge against each unit of output. Up to a certain point, variable costs fall as output increases, because division of labor can be employed. But the prices at which the entrepreneur can buy the factors of production are assumed to remain constant.

When constant and increasing costs were explained, it was assumed that all the entrepreneurs had arrived at the point of lowest total unit cost. They were assumed to be in an equilibrium position. Cost increased

because the sources of raw materials became scarcer. Hence cost of production for all entrepreneurs increased. The result was a higher normal, or equilibrium, price.

If we are to find an explanation of decreasing cost, we must find it in a change of some of the conditions laid down at the beginning of this chapter. If we assume that there are no inventions and no new methods to disturb a given industry or all industries, the cost of production for a marginal producer in any industry, and therefore for the industry as a whole, may fall for three main reasons: (1) the factors of production employed in the industry may decline in price; (2) the factors may be employed with greater efficiency; and (3) new and better factors may be used, or an improved application of the factors may be discovered. A fall in the cost of the factors, of course, would cause the price of the good produced by the given industry to fall, even though the demand remained the same. Here we seek some reason why an increase of demand will cause price to fall. Any outside force that caused one or more of the factors to fall in price would probably affect the cost of production of agricultural products, which are usually produced at increasing cost, just as it would the prices of goods commonly regarded as being produced at increasing cost. Such a fall in the cost of the factors cannot, therefore, be the condition that distinguishes decreasing cost from increasing cost.

There remain two main causes of decreasing cost as we have defined it. As stated above in (2) and (3), the factors of production may be employed with greater efficiency, and the use of *better factors already known and in use in other industries* may be adapted to this industry. The former cause involves the fact that if any fixed factor is employed with variable amounts of one or more other agents or factors, average costs will decline until the point of lowest total unit cost has been reached.

The relationships just stated have been explained in Chapter V. But it has been shown in the earlier chapters of Book II that *if perfect competition is assumed, entrepreneurs must make the best combination if they are to remain permanently in business*. There is only one condition under which an entrepreneur can remain in business without making the best combination, and that is the absence of perfect competition. It is possible, for example, for the producer of any good to make a very inefficient combination of the variables with the fixed factor and still cover all costs and even make a profit if there are no rivals to force the price down. It is difficult, however, to imagine a producer who could establish an ironclad monopoly in most lines of production. Since we are not here dealing with purely political monopolies, we must, therefore, find a cause

outside the realm of institutional forces to explain such a monopoly. Yet such monopolies do exist. For example, a small relatively isolated city may be served by an electric power plant. The cost of generating and distributing power may be higher than in a larger city situated in the same geographical environment, and the per capita demand for current may be the same in both. This is a condition that has often been observed. The explanation of the difference is that until such a power plant has a market of a certain size to serve, it cannot make as good a combination of the agents of production as a plant serving a market of that size.

If the reasons behind this "advantage of size" are sought, it is found that a number of causes are at work. Because the larger city has within its borders a greater diversity of economic activity, the demand for power is likely to be more continuous than in the smaller city and, therefore, the use of the expensive equipment of the plant is more continuous and the interest charge on the investment is spread over a larger number of units of output. Certain machines must be employed in every plant, and they cannot always be obtained in sizes suitable to the demand for their use. Because public utilities must render service with as little interruption as possible, every plant must have spare equipment to take the place of other equipment undergoing repairs and overhauling. In the large city plant the ratio of this "breakdown reserve" to the total investment may be smaller than in the small city.

In every plant there must be certain specialized labor. In the large plant this labor may be kept fully occupied with the specialized tasks for which it was trained or has natural ability. In the smaller plant it may be necessary either to do without some types of specialized labor or to employ it on work for which a cheaper grade of labor might be used in the large plant. Finally, the large plant can buy coal and supplies cheaper than the small plant, because it can buy in larger quantities.

The phenomenon of decreasing costs is frequently encountered in railway transportation. Here similar causes are at work. A double-track railway can carry more than double the traffic, and carry it at a lower cost per unit than a single-track railway. But there may not be enough traffic to occupy fully even a single track between two points. If that is the case, other things being equal, the single-track line will continue to operate, and its customers will be compelled to pay higher prices for services than if traffic were denser.

The small electric plant and the single-track railroad will experience decreasing costs until the demand for their services has increased to the point where they can fully utilize their plants. The fixed agents of pro-

duction will not need to be expanded as rapidly as the variable agents in order to offer to the public an increased amount of service. Finally the demand may become so large that not only will a single-track railway be as fully and as efficiently utilized as is possible, but a four-track or six-track road may be fully utilized, and a number of electric plants in a region may be linked together and utilized by this large market. At some point, however, the further expansion in the size of the productive unit will cease to cause further economies.

In Fig. 26 are represented the total cost curves of a plant serving a locality in which the demand at the initial period of time is represented by D_1D_2 , and at succeeding periods by D_3D_4 , D_5D_6 , and D_7D_8 . When the demand is at D_1D_2 the best size of plant is one that can produce, if fully occupied, an output of OF . But the demand curve cuts the total cost curve TT' at K , which is short of the point of lowest total unit cost. Hence not more than OE units can be produced and sold at a price that will cover cost, and two firms cannot sell in this market, because then each would produce only half the amount OE , and in that case the cost of each would be much above EK . When the demand has expanded so that in a succeeding period the demand schedule is represented by the curve D_3D_4 , the plant will produce at lowest total unit cost. But when the demand expands to the position represented by D_5D_6 , the plant must be used beyond the point of lowest total unit cost, or some of the buyers who are willing to pay the price equal to the cost of production at the point of intersection of the demand curve and the total cost curve, TT' , must go unsatisfied.

It will now be possible for the concern operating the plant to build a larger plant or a second plant under the same management. The railway can now lay a second track. The larger plant or the double-track railway will have total costs represented by tt' . It will be noted that the demand curve D_5D_6 cuts tt' at the point where it crosses TT' . It is therefore a matter of indifference to the management serving the demand represented by D_5D_6 , whether it overutilizes the original plant or builds additional capacity. When the demand has advanced to the position D_7D_8 , however, the additional plant will be built.

It is assumed that after the plant has grown to such a size that OH units are produced at the point of full utilization, N , a further expansion of plant will bring with it diseconomies of large-scale production and a rise of cost. Yet until the demand has increased to the point where this market can accommodate two plants having a capacity of OH when producing at the point of lowest total unit cost, that is, until the demand

very small sizes, and if skilled labor is to be employed it must be kept busy at the tasks for which it has been trained.

The second cause of decreasing cost referred to is the adaptation of improvements *already known and in use in other industries*. For example, it must have been apparent to the early producers of automobiles that if they could secure enough sales and therefore enough volume, labor-saving devices already in use in other industries, such as the conveyor system of assembly, could be adapted to the industry. Such advantageous devices are not taken over as a rule without modification, and therefore the word "adapted" is used and not the word "adopted." As the demand for motor cars increased and plants grew in size, machines were adapted from other industries which could not have been used when the output of cars was small. Improved methods of management and specialized labor likewise were brought in from other industries and adapted to the manufacture of automobiles. This is not to say, of course, that inventions and new methods of management were not also introduced by the automobile industry.

The causes of decreasing cost include also a number of more or less intangible elements. Among these is the mechanical principle that large containers and large machines often cost less in terms of material and labor per unit of productive capacity than small ones. The same principle applies to some extent to large buildings.

After all the economies attaching to a large plant have been realized, there remain certain cost-reducing advantages that attach to the operations of a large firm owning several plants. Buying in large quantities is cheaper, within limits, than buying in small quantities. It takes no more time for clerks to record a large transaction than a small one, and it does not take proportionately more time for laborers to sort out and load a large order. Still another element included in this miscellaneous category is the economy in the use of research workers of all sorts. Suppose that a hundred thousand dollars have been expended in research to make an improvement, in the routing of work in a factory, in advertising methods, on a machine, or in record-keeping. If the concern has a volume of sales amounting to a million dollars annually, the cost of this research per unit of sales will be higher than if the concern had annual sales of two million dollars. When several factories producing the same type of commodity and located in different parts of the country are brought under one management, transportation costs are often reduced. Competing factories ship goods into territories more easily reached from other factories; this is not done to the same extent if there are fewer firms.

None of these causes of decreasing cost would have worked out their effects as they did had competition been effective. For example, if there had been no consumer preferences for one make of motor car rather than another, if the consumer had been as indifferent to the brand or name of motor cars as he is to the name of the maker of nails, the working out of the improvements would have been quite different. Because each potential or actual maker of cars knew that new methods would be introduced as the demand expanded, each would have cut price against the others in order to secure larger volume and lower costs. This would have created a tendency toward monopoly from the beginning. Only when the market had become large enough to enable two or more firms to produce the same grade of automobile in sufficient quantity to enable both to obtain the maximum advantages in size available in the known state of the arts could perfect competition have put in an appearance.

We must now hasten to make two reservations without which the preceding statement appears unreal. In the first place, there was a very great change in the arts while the motor car was developing from its early form, when it was a kind of toy for adults, to its later stage, when it became a very useful implement for production and for pleasure. The decline in the cost of producing a car was probably due more to invention (not adapted improvements) than to any of the causes of decreasing cost considered in the preceding paragraphs. In the second place, it must not be supposed that we here imply that the production of motor cars ever has been carried on under the rigid conditions of perfect competition. Today advertising, the sales promotion of the local dealers, and the introduction of the style element (which makes last year's car in the low-price field as outmoded as last year's garments) have all contributed to the creation of somewhat separate markets for each type of car. None of them is sold in a perfectly competitive market.

DECLINING COST IS NOT DECREASING COST

It sometimes happens that an industry may experience declining cost without any of the increases in efficiency previously mentioned and without any perceptible improvements in technique taking place within the industry. When this happens, it is usually the result of reduction of material costs or transportation costs; that is, factors outside the industry itself may cause the cost of production to fall. When constant cost was under discussion we saw that a level price trend did not necessarily mean "constant cost" in the sense in which it is usually employed in economics.

Neither does a downward trend of price over a period of time mean "decreasing cost" in the technical sense in which it has here been employed, that is, a decrease in cost that must inevitably follow from an expansion of demand. The causes of a declining price trend are numerous. The firms that supply raw materials, machines, and building materials may be operating at decreasing cost, or inventions may make these agents of production cheaper. Transportation charges may be reduced, owing to denser traffic or to improvements. Finally, there may be improvements within the industry itself. All these dynamic changes may affect costs within any industry, whether it is subject to constant, increasing, or decreasing cost. They are not, however, the *results* of an increase of demand, which is the matter we are here considering.

LIMITATIONS ON DECREASING COST

A few of the causes of true decreasing cost appear to operate without limit. The larger a firm becomes, the lower will be some of its costs. The economy in the use of research staff is one example. It costs no more to make an improvement for a firm turning out a million articles a year than for one that turns out only a thousand of the same sort. Savings in transportation appear to increase until a firm attains a very large size, although undoubtedly a limit has been approached when the firm can operate an entire transportation system consisting of lake steamers and a railroad, as does one of the large iron and steel concerns. Most of the economies of a large plant operate, however, within rather definite limits. After a certain size has been reached, larger machines no longer bring lower costs unless there is an improvement in the machine amounting to an invention. The division of labor finally is carried as far as it can be carried in the existing state of the arts. These are limitations on the size of the individual plant. When they are passed, the costs tend to rise instead of fall.

There are also limitations on the size of the firm which may operate a number of plants. It is obvious that the limitations attaching to the size of a single plant do not attach to the firm as such. The causes that operate to make a firm less efficient after it has passed a certain size are as follows: (1) The purely supervisory staff increases more rapidly, after a time, than does the output. (2) The relations between different departments and plants become more complex and more likely to involve misunderstandings. (3) The number of regimenting rules and regulations, the red tape, becomes greater and leads to increased cost and

to inefficiency. (4) Lack of a feeling of responsibility and a growth of evasion of responsibility also make their appearance in both public and private organizations when they increase in size beyond a certain point.

It must be remembered, however, that improvements in production are not confined to mechanical inventions. Improvements in management, in control (accounting, routing of work and orders, and defining the responsibility of subordinates), and in the selection and training of supervisory personnel tend to offset the forces that make larger firms inefficient, just as mechanical inventions tend to offset some of the forces that make large plants inefficient.

DECREASING COST AND MONOPOLY

It sometimes happens that the demand for any product in a limited market is so small that, in terms of Fig. 26, not even one firm can sell an output of OH and still cover costs of production. Let it be assumed that the collective-demand curve intersects tt' in Fig. 26 at its lowest point. So long as one producer remains alone in the field, he can charge a monopoly price. He will not produce OH units, but something less than that amount. Others may try to compete with him and thus may increase the total output of the service or commodity. But if they do, the costs of both competitors will be higher than for one firm if they supply OH units, and still higher if they produce less than that amount. In neither case can they cover costs. They can reduce costs only by expanding the output. But expansion carries the price below HN , and again both will fail to cover costs. Instances of this type of competition have been given in Chapter XII. The conclusion holds that when the volume of demand is insufficient to accommodate two or more firms, that is, two producing OH units, monopoly is the only stable arrangement.

But even monopoly fails to produce stability in many cases. Knowledge is not perfect, and the intruder into the monopolized field may not know what is in store for him if he initiates competition. If he does initiate it, a price war is usually inevitable, and in many instances both concerns make losses until the futility of competition has been demonstrated and a combination is arranged. Often, of course, the intruder believes that the reigning monopolist has become slipshod and slumberous and that he can easily be ousted. Frequently this is the case. Thus it appears that decreasing cost offers no sure protection for monopoly. It has not done so in the case of public utilities. But the deterioration of service attendant on competition and other circumstances of a less

tangible nature have combined to produce a public opinion favorable to monopoly in the field of such services as the telephone, street transportation by trolley, and the supplies of gas, water, and electric current. Hence such enterprises are usually granted the exclusive right to render the service in a city or region. This having been accomplished, the state then undertakes to regulate the prices charged by the monopoly that has been created by it.

There has been much discussion in recent years concerning the necessity for monopoly in certain types of public-utility service. Telephone service is commonly accepted as typically monopolistic, because of the inconvenience and cost to the users when two systems operate over the same territory. Water usually is supplied by the municipality or by a water district. The belief exists that public health is better safeguarded if the local government furnishes the supply than if it is left in private hands. This belief has led to the establishment of public monopolies in nearly all the larger cities of the country. The cost of the mains for distribution is so great a part of the cost of distributing gas that it is generally believed that the service can be rendered most efficiently by a monopoly. There are, however, several outstanding cases of competition between publicly owned and privately owned electrical supplies. There is, as yet, insufficient evidence to warrant a sure conclusion in respect to the necessity of monopoly to secure lowest total unit cost for the supply of electricity in densely populated areas.

If there were perfect knowledge and perfect mobility of the factors of production, competition of many firms could not exist in any industry subject to decreasing cost. No matter what the cause of decreasing cost, there is always the possibility that a firm can reduce costs by expanding output, and if its sales decrease its costs will rise. Hence two firms selling in the same market will always have higher costs than one firm. As has been shown in Chapter XII, instability of price, price wars, and ultimately combination will always result.

Is it permissible, then, to speak of an *industry* that operates under decreasing cost, implying, as is usually the case, that there are many firms competing for sales? The answer is that we may do so only if we also assume that all these firms operate in partially sheltered markets and not otherwise. In Chapter XII it was shown that markets may be sheltered by transportation cost, by franchises and patents, and by the whims of consumers, which enable a seller to build a clientele. But even this assumption leaves the market subject to monopolistic competition. Under

this assumption we may, then, speak of decreasing cost as affecting a large number of firms.

From the conclusion in the paragraph above it follows, however, that decreasing cost is characteristic of unstable conditions, whereas constant cost and increasing cost are not. An increase in demand in respect to agricultural products, in the absence of changes in the arts, usually brings about a new equilibrium price, which will be higher than the price that prevailed before demand increased. An increase of demand affecting many firms producing in partially sheltered markets and working under decreasing cost may bring about a fall in price, but it will not bring an equilibrium price so long as decreasing cost prevails. Neither can we demonstrate what will be the exact effect on price in any particular case, as we can when constant and increasing cost are operative. Whether a monopolist or a partial monopolist will pass on all or scarcely any part of a reduction in costs to the buyers depends on all the conditions affecting the configuration of the cost curves and the demand curve and upon how great a degree of monopoly he actually possesses. If we use our earlier illustration of the owner of a brickyard in a small town, these points will be clear. Suppose that the demand for brick increases considerably, and suppose that the owner of the yard is operating under decreasing cost. At first, no doubt, he will simply sell more bricks at the same price. His profits will increase because his costs will decline. If he believes that by reducing price his sales will be greatly increased, and if he believes also that his total unit costs will decline rapidly with the increase in sales, he probably will cut price. He certainly will do so if he can predict both the amount of sales and the reduction of cost, and if he is an economic man. He may even go further and use the leverage he has gained to invade the market of a (monopolistic) competitor in another market by selling at a lower price than he charges in his own sheltered market. This is called "dumping" when practised in a foreign market.

Let us now suppose that for some reason the costs of production of the seller in the sheltered market rise. Will he raise his price? That depends on what effect the drop in sales which follows from higher prices will have on his costs and on whether the increase in price will cause a competitor in another market to come into his market. Conversely, if his costs fall, owing to some outside cause, the owner of the yard in a sheltered market may pass some or none of the gains to the buyers. If, however, the rise or fall in costs is general and affects all the producers in all the markets that are near enough to each other to compete, we may expect

that prices will be affected in the same direction as in a perfectly competitive market. Higher costs will mean higher prices, and lower costs lower prices. But the shifting of new costs or of reductions in cost to the buyers will not take place with the certainty and usually to the same extent as it does when perfect competition prevails.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Suppose you found that from 1900 to 1913 the cost of commodity X, when produced in typical factories, did not vary appreciably. Would this prove that the commodity was produced at constant cost? Do you need to know some additional facts before you can answer this question? If so, what facts?

2. Suppose that only 80 per cent of the productive capacity of the factories producing a good have been in use. An increase in demand now enables these factories to run at full capacity. Should you expect the total unit expense of the product to increase, decrease, or remain constant? If you found that the total unit expense decreased, should you conclude that the industry was operating under decreasing cost? Might it be operating under constant cost?

3. It is sometimes said that goods whose principal expense is common labor are likely to be produced at constant cost. Why, or why not?

4. "No good is produced at constant cost. If demand falls off, the producers will cut down production, employ fewer laborers, and buy less raw material. As a result, wages and the price of raw material will fall. If demand rises, higher prices will be paid for labor and raw material. Costs are thus shown to rise when demand increases and to fall when demand declines." Criticize.

5. In the production of an agricultural crop the following events take place: (1) the demand rises; (2) poorer land is cultivated, and the total produce increases 10 per cent; (3) the cost of production on the poorer land is 15 per cent above the marginal cost existing before demand rose. Later an improvement reduces the cost of production on the better grades of land, and some of the poorer land is abandoned, while the product of the better land is increased. Would it be possible for the competitive price to rise at the same time that average cost stood still? If so, is the commodity produced at increasing cost?

6. When war demand increased the price of wheat, previously uncultivated areas in Montana and Canada were broken up and sown to wheat. Do you suppose that the cost of growing wheat on these areas was higher than it was on good lands in Kansas? Why? After 1920 the price of wheat receded to near the level of 1913, if change in the purchasing power of money is taken into account. Yet much of the area broken up during the war remained in use. Does this show that the marginal cost of growing wheat on these lands is not higher than on the best wheatland in Kansas?

7. Suppose that all the land in a country was of uniform quality in respect to soil fertility, climatic environment, and location. As the population increased, at what point would increasing cost of the products of the land set in? Would it ever manifest itself?

8. The surplus-control bill of 1927 (McNary-Haugen Bill), which was passed by Congress and vetoed by the President, provided for the stabilization of the prices of certain agricultural products, principally wheat, cotton, corn, and swine. It was the intent of the bill that when a "surplus" of any of these products occurred it should be taken off the market at a price determined by a Federal board and either stored or exported. It was supposed that such action would prevent overproduction from depressing price unduly.

a. Suppose that the price of any of the products should be prevented from falling to its competitive price in any year. How would the crop in the following year probably be affected?

b. Assume that the prevailing price of wheat has been \$1.40 a bushel during the past five years and that the marginal cost of production has been \$1.80. If the board then should raise the price to \$1.80, how do you think future production would be affected?

c. Suppose the board by mistake raised the price above the competitive marginal cost for a period of years, how would future supply be affected?

d. If the surplus was sold abroad, how would foreign prices be affected? the incomes of farmers in foreign countries?

e. Do you think it possible to prevent fluctuations in the prices of agricultural products by governmental price-fixing? by storing surplus from years in which crops are abundant?

9. "The manufacture of automobiles is a business of decreasing cost. Users of motor cars are not, therefore, affected by the fact that wages of workers in that industry have risen. Since the industry operates at decreasing cost, the managers can always find some new way of cutting other costs; the increased wages bill will be met by the industry itself." Is decreasing cost correctly defined in this statement? Is it likely that increased wages could always be met in this manner?

10. It is said that decreasing cost cannot be of very much importance in agriculture, because farming operations are usually scattered over a very much greater area than is a manufacturing plant employing the same number of men. It is also said that decreasing cost cannot prevail in agriculture, because the economies of large-scale production cannot be obtained. Do you agree with these statements? What is your own conclusion as to absence of decreasing cost in agriculture?

CHAPTER XV · Industrial Combination and Monopoly

In the preceding chapters we have considered the general principles of price-making under competition and under varying degrees of monopolization. In this chapter we shall consider some of the commoner types of combination among industrial concerns. In the next chapter legal monopolies that rest on franchises, which are usually called public utilities, will be described. It already has been shown that unless a concern possesses the sole source of the supply of a given good, or unless competition is forbidden by law, as in the case of franchises and patents, a complete monopoly is difficult to establish and still more difficult to maintain. With the exception of patent monopolies, legal sanction is usually denied industrial combinations that have as their objective a substantial restriction on competition. Hence we should suppose that such restrictions among manufacturers and commercial concerns would be difficult to maintain. This a priori conclusion we shall find borne out in the following discussion.

When industrial combinations that restricted competition became important enough to attract public attention in the United States, many of them happened to be organized in a particular way. The organizers made use of an established legal device known as the "trust," whereby the stock of several concerns was deposited with a "trustee," who managed the business of the corporations whose stock had been deposited with him. Hence the term "trust" came to stand for any organization that aimed, or was supposed by the public to aim, at monopoly.

Before beginning the description of the various types of trust that have made their appearance in this country since the Civil War, it may be well to clear up two preliminary matters. The movement away from relatively free competition, which became important enough to attract attention about 1870, has not been the result of any particular change in legal or other institutions. With certain exceptions, noted later, the laws of the various states have always been and still are "hostile to monopoly." The tendency away from competition is not primarily due to any new contrivance by which men have been able to outwit

the law. It has been the result of certain technical conditions which have arisen in certain industries during the past century and a half and which became so strong in the latter half of the nineteenth century that men were motivated to seek ways and means to restrict the competition confronting them. These conditions can only be enumerated in a brief way in this book: (1) the increasing importance of fixed capital, which made competition more and more dangerous for those who engaged in it, and at the same time made intrusion by new firms more difficult; (2) the increasing vulnerability of capital to competition as its products became more specialized; (3) the increasing power of firms to bind consumers to them by means of advertising; (4) the increased use of patented machines; (5) in many lines, the economies of large plants and of large-scale operations in general.

These have been the primary factors in the growth of trusts. But institutions also have aided in bringing about the development of restrictions on competition when technical conditions were favorable. It is worthy of notice that in England, where until recently free trade prevailed, the trust movement was far less developed than in Germany and the United States, where home industry has been protected by high tariffs for more than a half century. The tendency toward a concentration of industrial ownership and control has been both an effect and a cause of combination in the United States. Whether we should have had this concentration (which is represented among other forms by interlocking directorates, the same men serving on the boards of directors of a number of corporations) without the growth of combinations cannot be determined, but it is unquestionably true that the use of various devices for concentration of control has furthered the development of trusts.

In analyzing the trust movement we must distinguish between the legal or institutional form that the combination assumes and the means by which it defends its monopolistic or quasi-monopolistic position. A monopoly may be organized as a holding company (one corporation owning or controlling enough of the stock of several subsidiary corporations to enable its board of directors to control the business policy of the subsidiaries), but its position may be defended against the competition of other concerns by means of patents, by ownership of natural resources, or any other barrier to competition. It is sometimes assumed that if all the producers in any line could form an airtight combination that would prevent any of its members from cutting below the established price, a more or less permanent monopoly would result. But

while such a trust doubtless would constitute a monopoly, its power would be only temporary unless it could also defend itself against the competition of new firms. The device by which it defends itself we shall call the "basis" of the monopoly or restriction, to distinguish it from the form that the organization takes.

THE POOL AND THE TRUST

One of the earliest forms of combination was the pool. The pool is an agreement among independent producers with a view to limiting competition. This may be done in a variety of ways. The profits pool, or earnings pool, is an arrangement by which all the profits, or earnings, are turned into a common fund and divided among the various producers according to a prearranged plan. The earnings pool was common in the railroad field in the seventies. Thus the trunk lines from Chicago to New York pooled their through-traffic earnings and allocated a definite share of the earnings to each railroad. The incentive to rate-cutting was thus withdrawn. Another form of pooling arrangement is the division of territory. Each company is given exclusive privileges in a certain territory, the other companies agreeing to keep out, thus eliminating competition. Another form is the agreement to limit output, each company being assigned a certain production per annum which may not be exceeded without penalty. Sometimes a central selling bureau is established, to which all the products are turned over and through which the sales are made. The central sales bureau, together with limitation of output, is the usual form of combination in Germany, and is known as a cartel (German, *Kartell*).

In the United States, pools flourished particularly in the seventies and eighties, though they have by no means been unknown since that time. Such contracts in restraint of trade and competition have never been recognized as valid in our courts. That is to say, if a member of a pool violated the agreement, he could not be sued for damages resulting to the other members because of his breach of contract, since the contract would not be sustained in court. Finally, legislation was passed, both state and Federal, which made these agreements not only invalid but even punishable by law.

Another form of combination resorted to in the eighties was the historical, or technical, trust. Originally the trust meant a special kind of combination. In this arrangement the controlling stockholders of the various companies entering into the combination turned their stock

certificates over to a board of trustees, who held the stock in trust for them as beneficiaries. Trust certificates were issued in exchange for the original stock certificates. In this manner the trustees secured control of a sufficient amount of the stock of each of the companies entering into the combination to enable them to control the various companies. They were thus able to elect the board of directors of each of the companies and so to secure a unified and harmonious management of the formerly competing concerns which now had entered into combination.

In the late eighties and early nineties, trusts (in the sense here used) were declared by the courts to be illegal combinations. A different form of combination was then resorted to—the holding company. The holding company is a corporation which owns, in its own name, the stock of subsidiary companies. Instead of a board of trustees holding legal title to the stock of various companies, we now have a separately organized corporation owning outright the controlling stock of the different companies. The holding company therefore is able to elect the boards of directors of the subsidiary companies and thus control the policy of all.

Finally, we have the merger, or consolidation. Here we have a number of separate corporations dissolving their separate identities and uniting in a single corporation, which takes over the property of the several companies. Unlike the pool, trust, or holding company, the companies combining in a merger, or consolidation, lose their identity as independently existing corporations.

THE SHERMAN ANTITRUST LAW

The courts refused to support the pooling agreements, and the technical trust likewise was found to be legally unenforceable. But other forms of combination, which will be described in later sections, were springing up. Legislation by the states proved of little use because there were always certain states that refused to pass antitrust laws. Then resort was had to Federal legislation. Because this legislation influenced profoundly the form that combinations took after 1890, it is necessary to discuss the early Federal legislation before describing the types of combination that succeeded the pool and the technical trust.

In 1890 Congress passed the Sherman Antitrust Law. The essential features of this law are contained in two sections. Section 1 states that every contract, combination, or conspiracy in restraint of trade among the several states or with foreign nations is declared illegal; section 2 declares that every person who shall monopolize or combine

or conspire with persons to monopolize any part of the trade or commerce among the several states or with foreign nations shall be deemed guilty of a misdemeanor. It is clear that the words "contract, combination, or conspiracy in restraint of trade" cover pools, trusts, holding companies, mergers, interlocking directorates, etc. The second section is broader. It includes not merely combinations but single corporations or individuals who seek to monopolize the market.

One of the earliest cases under the Sherman Law was the Sugar Trust case decided by the Supreme Court in 1895.¹ The government was seeking to dissolve the combination. The case involved the acquisition of a number of concerns manufacturing sugar in the state of Pennsylvania. The court held that while it might be admitted that this combination was seeking to monopolize the *manufacture* of sugar, no adequate proof had been presented in the case to show that the combination constituted a restraint of interstate *trade* or *commerce*. The combination therefore was not dissolved. It appears that the case was lost largely because the government failed to establish the direct effect of the combination on interstate trade.

Later (1899), in the Addyston Pipe and Steel Company² case, it was decided that this combination was illegal, because the various manufacturing companies had entered into contracts which directly restrained interstate commerce. Thus, at last, manufacturing combinations were held to come within the scope of the law.

Another leading case was the Debs case,³ decided in 1895. A strike had been called involving various railroad systems centering in Chicago. Injunctions were issued against the strikers as conspirators in restraint of interstate trade. Thus the Sherman Law was made to apply to labor unions, although the great manufacturing trusts, which Congress particularly intended to reach, at first escaped.

THE "RULE OF REASON"

Two important railroad cases were decided in 1897 and 1898 by the Supreme Court. They were the Trans-Missouri Freight Association⁴ and the Joint Traffic Association⁵ cases. In both cases the court held that these rate agreements were illegal under the Sherman Law. The court was divided, however. A majority of the judges held that the

¹ 156 U. S. 1.

² 175 U. S. 211.

³ 158 U. S. 564.

⁴ 166 U. S. 290.

⁵ 171 U. S. 505.

agreements were illegal, since the first section of the Sherman Law declared *every* contract or combination in restraint of trade to be illegal. The minority contended that well-established legal usage of the terms in question did not warrant this construction. They contended that under the common law it had never been held that *every* contract in restraint of trade was *ipso facto* contrary to public policy. Only contracts which unreasonably restrained trade were invalid. For example, a man may sell his store under contract not to engage in the mercantile business in that city for five years. Such a contract obviously restrains trade, but it is a reasonable restraint on trade and not contrary to public policy. In the opinion of the minority judges, the first section of the Sherman Act was to be interpreted in the light of the second section: the second section emphasized the intent to *monopolize* as the criterion of illegality. The minority judges contended that it was not the intent of Congress to declare *every* contract or combination in restraint of trade illegal, but only such contracts and combinations as aimed at monopoly—unreasonable contracts and combinations in restraint of trade.

This division in the court continued to appear in several decisions extending over a decade. Finally, in 1911, in the American Tobacco¹ and Standard Oil² cases, the personnel of the court had changed sufficiently so that the former minority construction now became the ruling of the court. These two combinations were dissolved, not on the ground that *every* contract or combination was illegal under the Sherman Act, but on the ground that evidence showed that they were consciously and deliberately *aiming at monopoly*. The combinations therefore constituted an unreasonable restraint of trade. Moreover, the court held that the certain proof of intent to monopolize was to be found in the long list of unfair practices by means of which these combinations had sought improperly and illegally to kill off competition. Such unfair practices are indubitable proof of intent to monopolize.

Thus the so-called "rule of reason" came finally to be applied to the Sherman Law. The law was interpreted to apply only to those contracts and combinations which are unreasonable; that is, those that had as their objective the establishment of a monopoly. The use of unfair practices was held to be a clear proof of intent to monopolize.

¹ 221 U. S. 106.

² 221 U. S. 1.

EFFICIENCY VERSUS UNFAIR PRACTICES

Let us suppose that a combination did not use unfair practices, but was able to attain a dominant place in the industry through efficient methods, and that the competitors were finally pushed to the wall, not by unfair practices, but through the superior efficiency of the combination. Such a combination might attain a monopoly through its superior efficiency in competition, without aiming or intending to monopolize the field. Would such a combination be guilty of unreasonable restraint of trade?

In 1920 the United States Supreme Court gave a partial answer to this question in the decision in the United States Steel Corporation¹ case. Here was a holding company which, at the inception of the combination in 1901, controlled about 60 per cent of the steel industry. As time went on, however, its control decreased to little more than 50 per cent. This combination controlled at least a half of this entire industry, but it was not using unfair methods against its competitors and was not, according to the findings of the court, pursuing a price policy contrary to public welfare. The court therefore refused to sustain the plea for dissolution. This position has been maintained in subsequent cases as late as 1931.² The practical effect of the decision is to permit combination, even though the combination assumes huge proportions, provided, first, that unfair methods are not used against competitors, and, second, that no public interest would be served by dissolution. It would seem, therefore, that we are tending in the direction of the German law with reference to combination. In Germany the cartel agreements are valid contracts enforceable at law. But this does not mean that a cartel may exploit the public by means of an extortionate price policy without being subject to prosecution in the courts. It is all a question of reasonableness. And it is this test of reasonableness that the Supreme Court has applied in the Tobacco, Standard Oil, and Steel Corporation cases, and others like them.

It should not be inferred, however, that the courts of the United States are willing to go as far as the German courts in permitting reasonable combinations. In the United States the courts are still skeptical of combination and strongly inclined to maintain competition wherever it is possible to do so. Pooling and trust agreements are still illegal in this country. Even certain "open price" associations have not always been

¹ 251 U. S. 417.

² U. S. v. International Harvester Co., 274 U. S. 693 (1927), and U. S. v. Swift and Co., 286 U. S. 106 (1931).

able to satisfy the courts as to the reasonableness of their activities. If, however, combination has been reached apparently as a result of natural economic tendencies without the use of unfair methods of attack upon competitors; if the combination appears to be economically efficient, so that in all probability the cost of production and the normal price would be higher were it broken up into less efficient competing units, then the courts will take the position that the test of reasonableness should be applied, and the combination should not be dissolved as a matter of course without regard to questions of efficiency and expediency. It is here that the rule of reason comes in to modify the former arbitrary and mechanical construction of the Sherman Law.

PATENTS

The clearest possible case of monopoly control is found where one concern continues to supply all the good that is sold. The simplest *method* by which this control can be obtained is through the ownership of patents. The difficulties with a monopoly based on patents are two: (1) the life of patents is limited, and (2) substitutes for the patented machines or consumers' goods may be invented. Hence this device for restricting competition, although simple, is not easy to apply.

An outstanding example of a partial monopoly based on patents is the United Shoe Machinery Company, which was formed in 1899 by a consolidation of several concerns that owned the basic patents covering many of the machines used in producing shoes. The *form* of the combination was that of the holding company—the United Shoe Machinery Company bought the stock of the other concerns—plus outright consolidation. From 1899 on, the stock or the property of many minor concerns was acquired. Here we have a combination, one of the principal bases of which was and still is the holding of patents essential to the production of a consumers' good. Still another basis was devised to make the combination more effective. Manufacturers of shoes were not permitted to buy outright some of the essential machines. Instead, the United Shoe Machinery Company leased this equipment to the manufacturers. Then it was provided further that unless the shoe-manufacturer used also other machines sold by the company, either the essential machines could not be used or a higher rental (royalty) had to be paid. This was the so-called "tying lease." The effect was to induce the manufacturers of shoes to take the machines of the company and not those of competitors.

It has been alleged that this combination was conducive to efficiency.¹ It provided shoe-manufacturers with the most up-to-date equipment as soon as it had been perfected, and it assured them that none of their competitors would be able to steal a march by introducing improved cost-reducing equipment of which they were ignorant. It has been asserted also that from the combination effected by the United Shoe Machinery Company great savings in production resulted.

The leasing system and the "tying" of leases were defended by the company on the ground that the owner of a patent may permit his improvement to be used as he sees fit and forbid its use where he desires to do so. When the government brought suit against the United Shoe Machinery Company the Supreme Court upheld the contention of the company and found that neither the leasing system nor the "tying" clause in the leases was in violation of the Sherman Antitrust Act.

PRICE LEADERSHIP AS THE BASIS OF CONTROL

If, in any industry, the smallest plant that can produce economically is very large relative to the total market, pure competition cannot exist. One of the best examples of this limiting condition is the steel industry. Before 1898 the steel industry had been more or less competitive. There were several large producers of half-manufactured iron and steel and many more fabricators. During the years before that date the relations between these producers may be characterized as an alternation of combination and competition. The pool and the informal understanding were the organization forms employed to restrict competition. But these agreements were never permanent, because there was no method by which the parties to the agreements could be compelled to keep their promises.

After about 1898, consolidations resulted in reducing the number of competing concerns. But these combinations did not greatly reduce competition, because they were frequently consolidations of the producers of the crude products, such as pig iron and steel forms, with the fabricators of these products, such as the makers of tubes, rods, nails, and bridges. In 1901 the United States Steel Corporation was formed. This concern as it finally developed was a combination not only of plants

¹This was the opinion of Justice Holmes of the U. S. Supreme Court, in *U. S. v. Winlow*, 227 U. S. 202, at 217. See Henry R. Seager and Charles A. Gulick, Jr., *Trust and Corporation Problems*, pp. 280-303, which should be consulted for a more complete treatment of the subject matter of this chapter.

producing iron and steel but also of mills producing tubes and pipe, tin-plate mills, coal mines, coking plants, iron mines, railroads, and lake steamship lines. It was an integration as well as a horizontal combination.

In 1901, about the time of its organization, the corporation controlled 43, 50, and 66 per cent respectively of the three leading products, pig iron, finished rolled products, and steel castings and ingots. The corporation, therefore, was far from being a complete monopoly through ownership. There was still a great deal of potential competition on the outside. Why the entire industry was not brought under the headship of the corporation cannot be stated with certainty. Fear of the law, difficulties of securing the stock of other concerns except at exorbitant prices, fear of unwieldiness, and belief that control could be established without complete ownership were probably some of the reasons. Some basis other than complete ownership had to be found.

The devices invented for controlling competition have been various. In the first place, the corporation introduced a policy of stabilizing prices for a time. It was believed that if the leader refrained from price-cutting, other concerns would be glad to co-operate. In the second place, the "basing-point system" of setting prices was introduced by the corporation. Many types of steel products, rails excluded, were sold from any of its plants at the price prevailing at the Pittsburgh plants plus the freight to destination from Pittsburgh. Thus, if a user of steel in Chicago sent his own truck to a steel mill in Gary, Indiana, and hauled away the steel he bought from that mill, he still paid the Pittsburgh price plus the freight from that point to Chicago. This system tended to make prices uniform and to reduce the opportunities for competition.

In the third place, after 1901 the "gentlemen's agreement" was introduced. The pooling arrangements that had existed before 1901 had been canceled by the corporation's subsidiaries. For these more formal arrangements, conferences between the officers of the corporation's subsidiaries and the officers of other producers were substituted. The "Gary dinners," to which were invited representatives of the other producers, also were used to persuade these producers to pursue a policy of co-operation. At these meetings the price policy of the corporation for the immediate future was explained and the other producers were urged to conform. It was admitted by the then president of the corporation, Judge Gary, that the objective of these conferences was to substitute "co-operation" for competition. There is ample evidence to show that, despite a decline in the share of the corporation in

the national output, it dominated the market for many years. Here we have an example of what has since come to be known as "price leadership." Other examples of leadership are found in the Standard Oil Company over the same period, the canned-salmon industry (Alaska Packers' Association), and the manufacture of crackers.¹

The description just given of the manner in which prices are influenced when price leadership prevails is scarcely satisfactory. Men do not usually forgo profits simply because the head of the largest member of the industry preaches co-operation. Ordinarily, they must be convinced that they will make more profits by following the leader than by refusing to do so. They can be convinced of the profitability of following in a number of ways. They may be frightened by having kept constantly before their eyes the losses that occurred during an earlier period when competition was unhampered, and by being shown what is currently taking place in other industries where co-operation is not employed. They can be influenced by fear of ruinous competition on the part of the leader if his prices are not conformed to. They can be coerced by threats to their credit if there is a close affiliation between the leader and powerful banks. Moreover, they can be kept in line, at least in periods of prosperity, if the leader sets the price high enough to permit his "followers" to make a comfortable margin of profit. Ordinarily, it is difficult, if not impossible, to discover, or at least to prove objectively, what have been the methods by which the leader has induced or compelled conformity to the prices he sets. This is necessarily so, because we are dealing here with the methods by which men's decisions are arrived at. "Fear of price-cutting" may be a very forcible factor in determining a seller's action, but it is impossible to measure it objectively.

It would be very instructive if we could discover how much the price leadership of the Steel Corporation raised prices above the competitive level. This cannot be done, however, because there is no adequate basis for comparison. The very fact that a successful combination brings about uniform prices removes the best standard we have for making the necessary comparison. This formidable difficulty stands in the way of all government regulation of monopolies.

An indirect method of judging the effect of limitation on prices is to compare the profits of the members of the combine with those "ordinarily" made in competitive business. Using this test, competent in-

¹ See Arthur R. Burns, *The Decline of Competition*, chap. iii, for an excellent discussion of price leadership.

investigators have asserted that the earnings of the corporation between 1901 and 1906 were in excess of those usually received by competitive manufacturing concerns.¹ It may be granted that this statement is correct, and yet it does not follow that the combine raised prices. It can be argued plausibly that these profits arose from the superior efficiency of the concern receiving them. This argument carries great weight unless it can be shown that even the least efficient followers also made more than usual profits, which would not happen consistently in a competitive industry. But if the least efficient followers did not make unusual profits, that would not prove that there was no raising of prices. The volume of business that the least efficient follower can secure may be so small that his costs remain high. To this it may be replied that this fact proves the superior efficiency of the leader. This is not a correct answer. The follower may not be free to secure additional sales by cutting prices. It might be true also that the least efficient producer was so inefficient that under competitive conditions prices would have fallen so low that he would have been forced out of business.

Still another indirect attack on the problem is to find out if the prices of the goods supposed to have been monopolized have exhibited greater stability after than before the combination came into operation. In the case of steel rails, the stabilization of price after 1902 is apparent, but for other important products the price was less stable.

A final argument to prove that monopoly prices were charged rests on the method of price-making. It is asserted that the Gary dinners, the basing-point system (which was subsequently dropped when it was attacked before the Federal Trade Commission), and the substantially simultaneous changes in prices by both the leader and the followers are persuasive to the effect that the objective of the corporation was to maintain prices above the competitive level.

It must be admitted that these are certainly not the phenomena ordinarily encountered in a perfectly competitive market. Yet they do not prove that prices have been higher on the average than they would have been under pure competition. Controlled prices may be higher than uncontrolled during certain periods and lower during other periods, but the average may be the same. Severe competition among concerns using much fixed capital in the past has led to bankruptcy of firms and the destruction of capital during periods of receding demand. These periods then have been followed by rising demand, inadequate supply,

¹Seager and Gulick, *op. cit.* p. 237.

and very high prices. It may well be argued that iron and steel have cost consumers no more on the average since 1901 than they would have cost if a greater amount of competition had prevailed.

Charges that the corporation and its subsidiaries existed in violation of the Sherman Antitrust Act were made almost as soon as it was organized. The government declined to take action, however, for about a decade. Finally, action to dissolve the corporation was brought and in the lower courts the corporation won. When the case finally came before the Supreme Court, that court held that violation of the law had not been proved.

SHARING THE MARKET

The pool has ceased to be a common form of combination, but the device of parceling out the market did not disappear when the pool was abandoned. The Standard Oil Company early made use of a selling agency in its attempts to control the market for refined oil. This agency was a corporation that bought all the crude oil, sold all the refined oil, and arranged the transportation of oil for all its members. It also determined the amount each of its members, all of whom were refiners, might sell. After 1890 this type of control was thought to be illegal and it ceased to exist openly. But in 1932 a group of soft-coal producers, who controlled only a small part of the national output, formed a selling association. This association solicited orders for its members, allocated the orders among them and set the sale price. But it did not determine how much each member might produce or the price at which he might sell outside the association. The activities of this organization were declared legal by the courts in 1933.¹

During the interim between 1900 and 1933, activities of associations for sharing the market were in abeyance, because of their apparent illegality. Yet there is evidence that by some means the same percentages of total production were maintained by ostensibly rival firms in several lines, such as meat-packing and the production of anthracite coal.

The object of the quota system, or sharing the market, is to remove the incentive for reducing the price of the product or increasing the price of the raw materials. If a firm is prevented from gaining additional business by a reduction of price, it will not cut price so long as demand remains constant. But a severe decline in demand tends to disrupt the quotas and to bring in price-cutting. Moreover, all quota

¹ 288 U. S. 344 (1933).

systems are in danger of being forced into price-cutting through the intrusion of new firms. It is evident, therefore, that some defense against competition from new plants is necessary if the scheme is to succeed. The best defense is a legal monopoly, but that is usually impossible. Another defense is the ownership by the group in the combination of the best sources of raw materials, such as the anthracite coal mines. But ordinarily this is out of the question. Still another defense is a large capital requirement for the establishment of a plant. Fear of the size and power of the group already in control is another deterrent. Several methods of getting rid of new competitors have been available—ruinous price-cutting, control of credit, and, in the past, secret rebates from railroads.

THE TRADE ASSOCIATION

The trade association is an organization of firms producing the same or similar goods. Sometimes these associations go by the name of "institutes." They exist in a variety of industries, ranging from the relatively unimportant lines of production to the steel and electric light and power industries. These associations have numerous functions. Information relative to methods of selling, accounting, and technical matters are exchanged among the members through the institute. Legislation believed by the members to be hostile and unfair to their business is opposed before Congress and the state legislatures. All these activities are of the type that any group of persons following the same calling commonly engages in.

In certain instances, however, the trade association was organized for a different purpose. The pool, the "one big corporation," and the various types of price agreements were tried successively, but great uncertainty as to their legality existed after 1890 when the Sherman Antitrust Act became law. The principal objections to all these types of restriction were that they had as their objective an enhancement of price above a fair level, and the fact that they were designed to prevent competition. The organizers of trade associations sought a type of co-operation among firms that avoided the appearance of a combination to restrain competition and at the same time discouraged "ruinous competition." Price wars in particular and price competition in general were to be avoided. Efficient producers were to be allowed and encouraged to charge prices that would permit a reasonable profit over and above the prime or direct costs of doing business.

The associations that desired to bring about a reduction of competition were forced to take into account the laws against combinations in restraint of trade. To attain this objective the associations collected and distributed certain kinds of information among their membership. The first step was to suggest to the member firms that standardized systems of cost accounting should be adopted. Then the production costs of the firms were collected, and average costs were computed and published among the membership. These costs did not include the normal rate of profit. Hence committees of the association sometimes determined the percentage of profit that should be added to insure a fair price. The circulation of these computations among the membership at least suggested the prices to be charged. But there is little evidence that anything more forceful than persuasion and verbal abuse was used against those who failed to conform to the "selling value" set by the association. In general it may be said that in times of brisk demand for their products the associations maintained conformity to their suggestions. But when depression, changes in demand, or overbuilding caused orders to decline, the membership usually failed to abide by the suggestions. "Live and let live" was followed as a precept as long as all could live, but it was very often abandoned by a member when his own survival was threatened.

Some of the associations collected and distributed additional information, such as the statistics of capacity, current production, unfilled orders, and inventories. That such information might well lead, and in many instances did lead, to a more economical ordering of production cannot be doubted. There can also be little doubt that it restricted competition to some extent and thereby kept in existence inefficient producers.

THE FEDERAL TRADE COMMISSION ACT AND THE CLAYTON ACT

After the Supreme Court had established the rule of reason it became apparent that the court alone could decide whether any given combination was violating the Sherman Antitrust Act. It appears to have been the belief of those who advocated and later defended that act that a few prosecutions would soon cause the large combinations to break up into groups of smaller competing concerns. But when it became apparent that only after long and expensive litigation could it be determined whether a given combination was acting in violation of the law, a demand for more stringent legislation arose. A further difficulty

grew out of the dissolution of concerns held in violation of the law. Moreover, it came to be widely held that large concerns were often more efficient than small concerns and that to break up the large enterprises into smaller units would reduce efficiency. Hence mere bigness ceased to be the principal criterion of intent to restrict competition. Interlocking directorates and "unfair" modes of competition now became the procedures against which the attack was directed. After several years of agitation Congress passed, in 1914, the Federal Trade Commission Act. This statute created a quasi-judicial commission upon which was conferred the power of exposing and stopping methods of unfair competition. The Clayton Act, which became a law a few weeks after the Federal Trade Commission Act, was designed to redefine "unfair competition." It replaced the Sherman Act to some extent. Its principal provisions forbid price discriminations in interstate commerce where the effect of such discriminations may be to lessen competition substantially and to create a monopoly. But mere size was no longer regarded as evidence of intent to mulct the consumers. Large concerns were to be allowed to do business so long as they did not seek to use their power to raise prices, and particularly so long as they did not attain their commanding positions by means of unfair trade practices.

From the very beginning the commission found difficulty in obtaining the facts upon which to bring actions or make the sort of report that would show whether the two acts of 1914 were being violated. The concerns into whose business it sought to inquire resisted the inquiry and appealed to the courts. The power of the commission was reduced by court decisions that limited its access to the facts. Through trade conferences it did induce manufacturing and selling organizations to adopt standards of fair competition, but the enforcement of its decisions still rests ultimately with the courts. Here again the tendency of the courts' decisions was to restrict the power of the commission and to make the courts the final authority as to what constituted "unfair" practices.

RECENT TRUST POLICY

As late as 1937 the national policy in respect to combinations was not clear. In the cases decided between 1926 and 1933 it appeared that the courts were still unable to identify mere size with monopoly. Yet it is correct to say that the national policy was one of hostility toward combinations. Exceptions to this policy were made from time to time. In 1918 the Webb-Pomerene Act exempted sellers engaged in exporting

from the Sherman Act and a portion of the Clayton Act. Agricultural co-operative selling agencies also were exempted by national legislation for the encouragement of these organizations, unless the Secretary of Agriculture should rule that the price of any product was unduly enhanced.

The exploitation of oil and natural gas gave rise to a different problem. Here, owing to geologic conditions, the owner of any limited surface area is often forced to exploit the resources under his land because otherwise adjacent well-owners will drain off the oil and gas deposited there by nature. When a field has once been opened it tends to be drained rapidly regardless of costs or prices. Several states have passed laws permitting combination for the purpose of controlling this waste. These laws have been upheld.¹ But the grant of power to the President to control interstate shipments for the same purpose was held to be an unconstitutional delegation of power. It appears, however, that Congress might control such commerce for the purpose of conserving resources.

The National Industrial Recovery Act passed by Congress in 1933 permitted the formation of codes of fair competition for the regulation of individual industries. Many of the codes, which were drawn up by representatives of the industry but required the approval of the President, imposed important limitations on competition. To be sure, these limitations were to be so designed as not to create monopolies but only to enforce fair competition. It is clear, however, that much of the unfair competition complained of was price-cutting. Some individual codes that followed the model code drawn up by the National Association of Manufacturers provided for complete control of output, but the administration refused to permit these provisions to go into effect except in the codes for lumber, copper, and petroleum, though other restrictions on competition were approved. The number of hours plants might be operated were restricted in about 60 codes. Many of the codes contained minimum-price provisions. Most of the codes that gave power over prices to the code authority prohibited sales below cost of production—a provision that was found nearly impossible of administration. Minimums could be set by the code authority also if prices were so low as to create “an emergency.” The codes frequently contained a clause that required sellers to publish price lists and adhere to them until they had been formally withdrawn.

¹ 284 U. S. 8 (1931), and 286 U. S. 210 (1932).

That there was potentially an enormous power to restrict competition in the NRA codes cannot be doubted. How effective such restriction became during the short time the law was in effect cannot be determined. It was declared unconstitutional by a unanimous decision of the Supreme Court in 1935.¹

In 1936 Congress passed an amendment to the Clayton Act.² In this amendment the actions that constitute unfair practices are set out in more detail than in the original act. It might even be said that Congress has attempted to substitute its own definition for those that have been set up or that might in the future be set up by the courts. These detailed definitions, like the provisions of the Clayton Act and the Federal Trade Commission Act, are aimed at unfair trade practices. But it is apparent that during the years that have elapsed since the passage of those laws public opinion has undergone a change. Then the principal fear was that a monopoly would be created when the concern that practiced unfair competition had driven all the "little fellows" out of business. The object was to protect the consumer. Secondly, these laws sought to protect the small producer. In the Robinson-Patman Act of 1936 the intent to protect the small man appears to be stronger, and at the same time another element makes its appearance. This is the intent to prevent chaotic competition, which may ruin all those who participate in it.

The most significant prohibition in the new law is one that forbids any person engaged in interstate commerce to discriminate between his customers in respect to price, service, or facilities when such discrimination may substantially reduce competition not only between him and other sellers on the same plane (two manufacturers or two wholesalers) but also between his customers. Such discrimination, however, may be justified by three conditions: (1) differences in the costs of manufacturing, selling, or delivery; (2) changes in price in order to meet competition; (3) a miscellaneous group of conditions including a general price change to meet market conditions, the sale of deteriorated or discontinued goods, sale under court process, or the equivalent of the foregoing.

Among the discriminations complained of was the so-called "unwarranted" quantity discount. Large retailers and especially chain stores were often able to purchase many sorts of goods considerably below the prices the small retailer was compelled to pay. This discount

¹ U. S. v. Schechter Poultry Corp., 55 Sup. Ct. 837 (1935).

² This amendment is commonly known as the Robinson-Patman Law.

had two effects: it enabled the recipient of the larger discount steadily to undersell the small man, and it enabled the manufacturer to charge what amounted to the same thing as a discriminating monopoly price. Of course, discrimination can be defended as good business policy on the grounds that the large retailers contribute to the stability of demand, that it costs less to serve them, and that there is always danger that they will establish their own brands if they cannot buy cheaply.

Another significant development of regulation since about 1930 is the attempt to legalize resale-price maintenance. Many years ago certain concerns that had created national markets for their branded or trade-marked commodities attempted to prevent retailers from using these commodities for purposes of competition. The manufacturer set the price at which the commodity was to sell at retail. But retailers often found it to their advantage to reduce the prices of these nationally known goods and advertise the reduction among their potential customers. It was advantageous for them to do this because the prices at which the goods ordinarily sold were known by a majority of people, and the statement that they were to be had at a reduction could easily be verified. As a result, such advertisement often brought many customers to the retailers who were in a position to make the cut in prices. But not all retailers were able to draw in custom. The small stores that could not afford advertising space in the large metropolitan newspapers and the stores that served merely neighborhood trade could not take advantage of this method of increasing the volume of sales. They lost business through the operations of the price-cutter. These stores complained to the manufacturer and threatened to refuse to sell the good. Moreover, the manufacturer who had created in the minds of the buying public the belief that his commodity was very superior and ought to sell for a higher price than the ordinary commodity—as a branded, advertised soap in comparison with “ordinary” or “common” soap—did not like to have his commodity sold at only a little more than the ordinary one. People might come to believe that it was worth no more and refuse to buy it at the high price.

To protect his outlets, the small retailers, and to protect the prestige of his commodity, the manufacturer refused to sell to dealers unless they would sign an agreement not to cut the price when they in turn sold the good. This is the resale-price maintenance agreement. But the courts refused to enforce these agreements when the first cases came before them early in this century, even when the commodity was a patented article. The manufacturers then resorted to refusal to sell to

dealers who consistently cut prices. But where the goods were distributed through wholesalers the manufacturer could not discover, without making expensive investigations, where the price-cutting firms obtained their supplies, and when one manufacturer did establish a systematic investigation his acts were held to be illegal.

Among the small retailers, especially in the drug trade, agitation for the legalization of price maintenance continued. Since 1933 a number of states have passed laws designed to limit certain competitive practices of the chain stores and the large metropolitan department stores. They are usually called "fair-trade" laws. In some of the states these laws make it illegal for any dealer to sell any commodity for less than the price stipulated in the contract entered into with the manufacturer. Congress has not yet passed such a law, but there is agitation for it, and the antitrust acts may be amended to legalize and make enforceable in respect to interstate commerce what is apparently valid for intrastate trade.

The chain stores and at least one large department store have attempted to counteract the effects of these laws by making a comparison between the prices of commodities for which the prices have been fixed and the prices of those for which no such practice is in force. The chains often have their own brands of goods, which are either produced in factories owned by them or are made by concerns that have no nationally advertised brands. Handbills, newspaper advertisements, posters on store counters, and labels on both classes of goods (those on which the price is fixed and those on which it is not) have been employed to call the attention of the buying public to the effect of price maintenance. It is possible that the attempt to prevent retailers from reducing the prices of nationally advertised brands will collapse. It is possible also that it will not, and that the chains and large metropolitan department stores will either manufacture to a greater extent than now the goods they sell or will enter into agreements with manufacturers to produce special brands for their use alone. In either case the markets for these types of goods will become further split up into more or less non-competing groups of commodities.

SUMMARY AND CONCLUSIONS

In the preceding discussion, we have seen that the tendency away from relatively unfettered competition has been going on steadily since about 1870. The tendency to combine first became important among

railroads and industries using much fixed capital, such as oil-refining and iron and steel. But as time went on it spread to other types of manufacture and in recent decades has become common in retailing. Since 1920 there have also developed in agriculture definite movements toward limitation of output through co-operation.

The causes of this drift away from competition and the recession of the spirit of free enterprise are not entirely clear. Among the heavy industries, and especially where the economies of large-scale production are marked, the reasons are fairly obvious. No doubt the rapid pace of invention and other improvements in recent years has had something to do with the widespread desire for security in contrast with the more adventurous spirit of business during the first part of the nineteenth century. In part, changes in the attitude of people toward the rights and functions of the individual in society have been a factor. As is usually the case, government has been more conservative than leaders of business and has tried with some success to stop the growth of combinations. But it is evident that before 1914 the government prevented the use of certain forms of organization while it did nothing to attack the bases on which the combinations rested. Even after the passage of the Federal Trade Commission Act and the Clayton Act, not very much was accomplished to prevent the use of certain of the more obvious means by which the combinations defended themselves against price-cutting within the combining group and against the competition of outsiders.

At the present time we are witnessing a wider and more comprehensive movement toward co-operation among the members of the same trade and occupation. Labor unions are more powerful than they have been at any time in the history of the country; co-operatives among agricultural producers and exporters have been fostered by special acts of Congress; and still another piece of legislation, the National Industrial Recovery Act, set up conditions that might have led to further limitations on competition in the industrial and commercial fields. This act was declared invalid, but the demand for some type of organization that shall eliminate "chiseling" is insistent and may well result in further enabling legislation. Whether it does or not, the spirit of business in the United States is different from that of fifty years ago. Further steps to place impediments in the way of unfettered competition may be expected.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Is it good public policy to permit concerns to expand beyond the point where they can produce at a minimum cost? What motives might lead to such an expansion? Why might competition not prevent such expansion?

2. Why has it been relatively easier for manufacturers to combine than retailers? than farmers?

3. What is meant by ruinous competition? Is all competition ruinous? Is it wasteful as compared with some other conceivable relationship between different producers? with which other relationship?

4. What is the purpose for which the national government grants patents? Suppose that all research workers and all other employees of a given concern sign contracts to turn over to the employer all inventions they may make during the term of their employment. Is the purpose defeated?

5. At one time the courts appeared to determine whether the intent of a combination was to restrict competition by the proportion of the national output controlled by the combination. Was this test economically valid? Would it be valid if the question was to determine not intent but actual control?

6. Discuss the following proposition: Unless not more than five large concerns are allowed to produce all of an important product, the maximum economies of large-scale production cannot be achieved; but if only five concerns produce they will almost certainly set prices so high that profits will be higher than those usually made by businesses of similar risk. This is monopoly. Hence, the five concerns must be broken up into many smaller concerns.

7. In some communities the physicians' societies publish and advise standard fees for certain types of medical service. The rich may be charged more and the poor may be attended for nominal fees. Is this a combination to restrain trade? Can you see any reason why physicians should be allowed to form an open-price association and the manufacturers of furniture not be allowed to do so?

8. Why was it desirable for the United States Steel Corporation to acquire iron mines? Suppose some other concern had obtained control of all the deposits of high-grade iron ore, how would that affect the possible profits of the steel concern?

9. If it was desirable for the Steel Corporation to acquire mines, why is it not thought desirable for the cotton mills to own cotton farms? retail stores?

CHAPTER XVI · Public Regulation of Prices



THE GROWING IMPORTANCE OF PUBLIC REGULATION

Beginning in the eighteenth century in England, and then spreading to other countries, there came a movement for the restriction of government interference with the economic activity of individuals. Before that time regulation of many sorts of individual enterprise had been common. At least it had been legal for the government to regulate, but, as is usually the case, the legal scope of the regulations exceeded the actual accomplishments of the public authorities, just as, at the present time, the legal restrictions on motor-car traffic far exceed the actual enforcements in most communities.

So far as present indications go, we are now again to experience a much more extensive application of government regulation of economic activity. The objective of some regulations, such as traffic ordinances, is chiefly to protect the physical safety of persons and to prevent damage to property. Others, such as laws protecting the customer against the sale of foods and drugs containing harmful substances, seek merely to safeguard public health. But the number of regulations the objective of which is to protect or to enhance the economic interests of large numbers of people is constantly growing. The mass of regulatory law that has emerged from the national Congress since 1933, commonly referred to as "New Deal legislation," is ample proof that at the present time the people of the United States are convinced that regulation of economic activity for the protection or enhancement of the economic interests of large groups should advance and not recede. In European countries the amount of regulation attempted by the state not only is greater than in the United States but also is constantly increasing. And the form of political organization appears not to cause much difference in the drift toward economic legislation. In socialistic Russia, in Fascist Italy, in Nazi Germany, and in the more democratic nations, such as England and the Scandinavian kingdoms, the story is generally the same: the amount of state regulation is constantly increasing.

In the United States, economic regulations have become complicated

and numerous. In general, however, they take the form of regulations of production and prices. The AAA legislation sought to control production and prices of agricultural commodities. The NRA sought to control production and prices in respect to all the groups that formed codes in manufacturing and trade.

There is little historical evidence that will lead us directly to conclusions about the effects of many of these newer regulatory measures. As has just been said, we in English-speaking countries have lived for more than a century under a system of social policy that left the individual relatively free (and because free, unprotected) to work out his own economic problems. It is probably almost futile to expect much light from the experiences of people with the more extensive (mercantilistic) regulations during the sixteenth and seventeenth centuries, because economic conditions then were very different from those now existent. There is, however, one type of regulation which persisted throughout the period of individualism and with which we have had a great deal of experience in the United States during the past three decades. When nearly all other types of production escaped from the guiding hand of government officers during the eighteenth and nineteenth centuries, a few were unable to free themselves, although even they secured a partial release. These were the "businesses affected with a public interest," or, as they are commonly known, *public utilities*. From our experience with regulation in respect to these businesses some light can be shed on the problem of regulating prices and production under a system of private property. Hence it is useful to examine briefly some of the problems arising in the attempt of the government to control these public utilities.

THE NATURE OF A PUBLIC UTILITY

The businesses that did not escape from government regulation and that have been regarded by legislatures and held by courts as peculiarly subject to governmental supervision are usually called public utilities. All business from time immemorial has been restrained by the laws of property and contract and by such laws as are required to protect the health and safety of the public. But public utilities have been peculiarly regulated. The charges they make for services sold to the public, the kind and quality of the service they sell, whom they must serve, financing of the operating companies, the records they must keep, and the wages they pay labor have all been regulated at one time or another for many

years. Private businesses sell at whatever prices they choose, sell to whom they please, sell the quality of wares they wish (so long as they do not collide with regulations protecting health and safety), raise capital as they see fit (so long as they do not defraud investors), and pay whatever wages they find necessary in bargaining with their employees.

The legislatures and the courts distinguished between public utilities and private businesses on several grounds. These grounds have been called the "characteristics" of public utilities. In the first place, to be a public utility a concern must supply some service that is essential to a civilized existence, but it need not be an absolute necessity as the term is usually understood. In the second place, the service must be one that is sold generally to the public. In the third place, there must be a tendency for the business to become a monopoly.

Only the second and third characteristics require explanation. The second characteristic is not entirely clear. Several examples may help to make the meaning somewhat more understandable. The business of supplying electric current to the public has been generally declared to be a public utility or a "business affected with a public interest." But if a manufacturer generates current for use within his own factory and then sells some of it to someone else, he will not ordinarily be declared to be engaged in operating a public utility. If the owner of a fleet of trucks serves only a limited number of customers, who arrange with him by private contract to haul for them, he is not ordinarily engaged in a business affected with a public interest. The rule is that the supplier of current or of transportation by truck must "hold himself out" to serve anyone who applies, before his business comes under public-utility regulation. It is evident that the reason for the distinction is to be found in the need for public protection in the one case and the absence of that need in the other. The transport concern that serves a large public or the supplier of electric current who serves many consumers usually has these consumers somewhat at his mercy. This is especially true if the electric company or the transport concern is the only one serving a certain territory.

The third characteristic, that of a tendency toward monopoly, has already been partly explained in the preceding chapters. There are certain types of business that can operate at lower costs (within limits) if they control a relatively large market. The railroad, the streetcar, the supply of gas, the telephone, the supply of water, the telegraph, and probably the supply of electric current are among these natural monopolies. Some of these services probably would be made monopolistic

by legislative action even though decreasing cost did not drive them toward monopoly. To have two telephone systems in a community is inconvenient. To have access to all other users each person must be a customer of both systems. More poles and wires are necessarily placed in the streets. This means needless duplication, inconvenience, and cost to consumers. Competing streetcar systems are not only unprofitable but also a needless obstacle to other street traffic. Competing gas companies are likely to be costly and at the same time to cause needless excavation in public highways. To some extent the public is inconvenienced also by competing electric light and power systems.

It has been said that these public utilities have a fourth characteristic, the right of eminent domain and the permission to use public highways. Most public utilities have these privileges, but the privileges do not make them businesses affected with a public interest. On the contrary, they have these privileges because they render an essential public service. They may take private property against the will of the owner, after paying a reasonable price for it, because no one owner of private property may be allowed to interfere for his own pleasure or profit with the economical operation of such a service. They are allowed to use public highways either because they better serve the public on that account or because the cost of the service to the public is thereby reduced.

THE REGULATION OF PUBLIC-UTILITY RATES

The principal reason for regulating public utilities is to control the prices they charge consumers for the services they sell, because the utilities are usually monopolies and without regulation would exploit the consumers. The prices have come to be called public-utility "rates." The legal doctrine, stated in broadest terms, is that such rates must be "reasonable." That is, they must be reasonable both from the point of view of the consuming public and from the point of view of the investors in public-utility securities. The factual definition of "reasonableness" has been a difficult problem for the courts and legislatures. It is generally agreed, however, that such rates are reasonable for both the consumer and the utility when they allow the investors in such properties to earn on the money they have invested the return they could have earned if they had invested in competitive businesses of comparable risk. Behind this abstract statement lies the assumption that competitive prices are reasonable prices. If the customer pays one seller no more for a service than other competitive sellers are asking he should not complain; if the

seller receives as much as his competitors are getting he has no ground for complaint.

The first step in determining a reasonable rate is to find the value of the property. This value is called by the courts the "fair value" of the property. It is referred to also as the "rate base." At first sight this might appear an easy step to take. The records of every utility concern, it might be supposed, would show the sums of money contributed to the capital of the company by the buyers of stocks and bonds. But it is now generally held by the courts that these contributions are insufficient evidence of the value of the investment. In the first place, the company may have reinvested the legitimate earnings of the concern, which otherwise might legally have been disbursed as dividends. If they have, the value of the investment of the stockholders has increased above their original contribution. In the second place, the property may have been allowed to deteriorate, and moneys may have been paid out to stockholders that should have gone for repairs and replacements. In the third place, the property may have risen in money value owing to a rise in the general price level or to increases in the costs of construction peculiar to utility property. In the fourth place, the records of the company may have been lost or they may have been kept in such a manner that it is impossible to discover what have been the actual contributions of the holders of securities.

MARKET VALUE OF UTILITY PROPERTY

In the past it has often happened that the first difficulty encountered by regulatory bodies has been the unavailability of records. An appraisal had to be made. What, then, should be the criterion of the fair value of the property? It has been suggested that the comparable procedure in appraising a competitive property is to find out what it would sell for, assuming willing buyers, or what is its market value. But the large properties of public-utility concerns are not bought and sold like cloth, land, and securities. Seldom is such a property sold at public auction. If a railway system is acquired by another company it is usually through the purchase of its stock. So it was suggested next that the market value of the stocks and bonds of the utilities should be taken as a criterion of the fair value. Here again practical difficulties were encountered. In recent years the stocks of many operating companies have been held by holding companies and therefore not offered for sale on the exchanges.

The fatal objection to this criterion which serves so well as guide in the appraisal of property used competitively is, however, the fact that the value of property depends on its earnings. A railroad that is prosperous is worth more as a going concern than a substantially similar property that earns less. The stock of the former is almost sure to be worth more. But earnings depend on two things: the prices charged for goods and services sold, and the operating expenses. Now the purpose of regulation is to determine whether the prices charged for the services of a utility are reasonable or not. Clearly this determination cannot be made by assuming that the market value of the regulated property is correct in the beginning, because that value depends partly on the prices charged. If the market value of the property were to be taken as a standard for determining fair value, regulation would be a futile proceeding in most cases.

REPRODUCTION COST

In 1898 a famous case, *Smyth v. Ames*, was decided by the United States Supreme Court.¹ It concerned the regulation of the rates of a number of railways by the state of Nebraska. The attorneys for the state argued that the true test of the fair value of the properties of these concerns was the cost to reproduce them at the time the regulation was imposed; the attorneys for the railway companies argued that what it had originally cost to produce them was a better test, although they contended that the par value of the stocks and bonds issued by the companies also should be considered. The court handed down a confusing decision, stating that a number of factors, among which original cost and reproduction stand out most prominently, must be taken into account. At the time this suit was begun the general price level had been falling for about a generation. Consequently the cost to reproduce the properties had been declining. Naturally the railways did not want their investment reduced in the same proportion, and quite as naturally the consumers could not see why the rates of the railways should remain high when nearly all other goods and services had fallen in price.

For many years after this decision the general price level rose, slowly until the outbreak of the World War and then very rapidly. Some of the properties valued during this period had been built at earlier and lower costs and some at later and higher costs. There was no definite

¹ 169 U. S. 466 (1898).

rule that the commissions could follow in valuing railroads and other utility property. Probably reproduction cost played a more important role than original cost.

But after the close of the war, when prices and costs showed no signs of retreating to the level of 1913, the utilities made a drive to have reproduction cost adopted as the only basis for valuation. The arguments of the attorneys for these companies appear to the lay mind unnecessarily complicated and tortuous. The gist of them amounts to this: In competitive businesses the value of fixed capital *that is still efficient* rises and falls with the rise and fall of the general price level. Since this is the rule for competitive business where the prices of products and services tend to agree with costs and where monopolistic exploitation does not take place, it appears just and reasonable that the rule should be applied to the fixed capital of such regulated natural monopolies as utilities and railways. This economic argument appears unanswerable.

Two counterarguments are commonly brought against it. The first rests not on economic reasoning but on grounds of social policy and on legal theory. It is said that all these regulated businesses are agents of the state and as such are not entitled to profits caused by rising prices. To this it is replied that if the utilities are merely the agents of the state, they should not be compelled to take risks such as competitive business encounters. If they are to be limited to a fair return on the original cost of their investments in times of rising prices, they must in good faith be guaranteed against loss caused by declining prices. The rule cannot be, "Heads I win, tails you lose."

The second counterargument rests on the denial of the administrative feasibility of the cost-of-reproduction rule. Granted that the rule is economically correct and ethically sound, it cannot be made to work. There are a number of reasons. In the first place, to make regulation effective it is necessary to make frequent valuations to keep the rate base currently correct. This is all the more necessary when the prices of materials and wages are constantly changing. But this requirement is impossible of performance. The valuation of a large property, such as a major railway system or a large electric light and power plant, is a tedious and costly procedure. The public has not been willing to appropriate the funds to make it possible to keep valuations up-to-date except in the case of the railways. Moreover, so time-consuming is the process that the valuations may be out of date when they are finished.

Another practical difficulty arises out of the progress of the arts. In

competitive business obsolete and obsolescent buildings and machinery are never worth what it would cost to reproduce them. Since their earning power has been reduced by the appearance of more commodious or efficient substitutes they are worth only the present value of what they can earn. Under competition the buyers and other users of these obsolescent properties set values on them by the higgling of the market; similar properties are always being bought and sold. But public-utility properties are seldom bought and sold. There is no jury of entrepreneurs who by their bids and offers determine what the properties are worth. Neither can a competitor enter the field and produce and offer the services for sale at a lower price because his costs are lower, owing to the use of newer processes and equipment. An automobile-manufacturer can no longer sell at the old price a car that does not include such recent improvements as a six-cylinder engine, a self-starter, four-wheel brakes, and a maximum speed of more than sixty miles an hour. But a utility can often compel consumers to use a type of equipment that is out of date and to accept service that could easily be improved. It was only after the private motor car and the autobus had cut their passenger business to a fraction of its former volume that the railroads found they could run trains that average sixty miles an hour for a trip of several hundred miles. But the adoption of improvements often necessitates the scrapping of old equipment. It is often more profitable to keep the old equipment. What is the old equipment worth? Here the argument for reproduction cost breaks down because the analogy with competitive business breaks down. The old apartment house or the old plow is not worth what it would cost to reproduce it, even after allowance for wear and tear (accrued physical depreciation). Nor by competitive standards are the old boiler and generator that can convert two pounds of coal into only a kilowatt-hour of electrical energy when improved equipment can convert that amount of fuel into one and four-tenths or even two kilowatt-hours.

When the engineer whose duty it is to value the equipment for purposes of regulation encounters this old, obsolescent equipment, he has no guide to follow. He cannot appraise the machinery according to its earnings, which would be one guide for estimating the value of a machine used in competitive industry, because the utility is a monopoly. He cannot accept cost to reproduce in existing physical condition, because the machine is inefficient by current standards of performance of newer machines. But even if he accepts cost to reproduce in existing condition, his difficulties are not entirely obviated. Some of the old equipment is

no longer produced, and much more has come to be made of different materials and by different processes. The best that he can do is to make an honest guess at its reproduction cost. But where guessing enters, accuracy flies out of the window. Hence arise controversies between valuation experts testifying before courts and commissions, endless arguments, delay, and, worst of all, hesitancy on the part of commissions to begin proceedings to regulate prices.

A further, but by no means final, difficulty is encountered in ascertaining reproduction cost. Not all the cost of building up a business as distinct from a physical plant is apparent to the eye of the appraiser as he explores the plant of an electric light and power concern or as he passes through the yards and over the line of a railway. Temporary structures that may have been necessary during the construction of the plant but were scrapped when it was completed have disappeared. Such intangibles as costs of organizing the working staff for offices and plant are not apparent to observation. The price that a small user of land would pay for a right of way or for a plant site is not necessarily the price that a utility would have to pay. Often the utility must use particular pieces of property; hence they are permitted to condemn land by legal procedures and to take possession of it after paying the owner a price set by a suit at law. The verdicts of juries are often erratic, and court proceedings are expensive. To avoid these uncertainties and costs, utilities often pay more than other willing purchasers would offer for land. The difficulties mentioned in this paragraph do not complete the list, but they should show how complicated is the problem of appraisal for rate regulation by the reproduction-cost standard.

Original cost, by contrast, can be ascertained with reasonable accuracy from the company's books if they have been kept properly. But it is unworkable in a changing society. Thus regulation today is in a dubious position. No one knows whether it can be made to work or not. The rule that appears to have most of the economic arguments in its favor, reproduction cost, is extremely difficult to administer, and the rule that can be easily administered is economically correct only when the costs of construction do not vary widely over measurably long periods of time.

THE RATE OF RETURN

After the fair value of the property of a utility has been settled, the next step is to find the percentage a company may be allowed to earn on this fair value. Here again we encounter a great deal of guess-

ing. Many years ago the Supreme Court of the United States decided in a particular case that 6 per cent was a fair return. Commissions and inferior courts copied this decision until rising interest rates during the period 1914 to 1920 made it necessary to adopt a higher figure. The higher Federal courts then said that from 7 to 8 per cent was a fair figure. The inferior courts and the commissions again followed suit. Today there is a tendency for courts to set lower rates of return.

It must be clear to every student that the rate a utility should be allowed to earn ought to vary with the conditions under which the business is conducted. Contractual interest rates and dividends on common stock are not solely payments to savers for forgoing the consumption of their funds. Both types of security, bonds and stocks, fall far short of the certainty of a government bond. Risk is involved in both. But the risk involved in lending to a corporation operating a telephone service in a densely populated district that is constantly expanding is quite different from that involved in lending to a street-railway company in a Middle Western town that has ceased to grow rapidly. The risk assumed by the stockholders also is different. Obviously the rate of return ought to vary with conditions affecting the risk imposed on the investors.

There is a variation also in the competitive rates of interest in different parts of the country. This in turn should give rise to variations in the rate of return. Moreover, the rate at which investors are willing to invest in the securities of any company depends in part on the confidence they have in the management of that company. If the past performance of the management warrants great confidence, the rate will be relatively low; if past performance is not regarded favorably by the investors, the rate will be relatively high. But the variations in neither risk nor money rates can be accurately ascertained. The best that can be done is to examine the financial history of the company and the rates of interest paid by other concerns and to take the testimony of bankers and investment houses. There is much leeway, however, and a commission or court may honestly miss widely the hypothetical fair rate.

OPERATING EXPENSES

The charges that a railway or a municipal utility may make against the public, according to the theory of regulation set up by the legislatures and courts, are made up of two groups. We have considered in the preceding section the charge for the use of investment, the fair return on the fair value. The second group includes the direct expenses

of production incurred in the current operation of the plant. These consist of costs of materials, fuel, repairs, wages of ordinary labor, salaries of the executive staff, and finally a charge for the depreciation of the fixed capital. It might be supposed that there could not exist in respect to these items any high degree of uncertainty. All but one of these payments (depreciation) are usually made to persons outside the company, to laborers, technical men, manufacturers, and owners of coal mines. For a long time the theory of the courts was that none of these payments, except depreciation, were subject to regulation by the commissions unless there was substantial proof of fraud.

Recent developments have shown that the position of the courts that these were competitively determined costs and therefore that no cause for overcharging could be concealed in them is not always warranted by the facts. Company A may buy supplies, equipment, or current from Company B, which is a subsidiary of the same holding company that owns all the stock of Company A or the controlling interest in it. There are instances in which engineering services or natural gas are bought by one subsidiary of a holding company from another subsidiary. Clearly the opportunity is open to the holding company to increase the charges of Company A to its consumers by increasing its expenses. On the face of the record, on the books of Company A there are no excessive earnings. But the concern that sells to Company A will make high profits, which will flow to the holding company. May not the state step in and regulate the prices paid by operating companies for equipment and supplies? It appears that legally they may do so, but practically the difficulties are always very great and sometimes almost insurmountable. Producers of equipment and owners of mines are not public utilities and not subject to regulation. They are often located in a different state from the operating utility company.

Somewhat the same difficulty arises in respect to labor. The operators of railway trains are organized into strong unions, and a large part of the cost of conducting transportation consists of the wages paid to these workers. If their wages rise considerably, there is a strong argument for raising freight and passenger rates. To be sure, this case differs from the one just discussed, because the investors in railway securities stand to lose rather than gain from the rise in wages. But if the rates of railways are to be limited to the same return that the agents of production would receive in competitive enterprises, the essential problem is the same. The laborers and executive staffs are no more entitled to profit by the monopoly position of utilities than the investors. High

charges for the services of these utilities can arise from excessive salaries and wages as well as from dividends on excessive valuations.

Probably the most difficult expense of production with which a regulatory commission has to deal is the charge for depreciation. The owners of the plant obviously are entitled to collect from consumers an amount represented by the wearing out of the fixed plant. A part of this expense is represented by repairs made at frequent intervals during the year. But large pieces of equipment all ultimately wear out in spite of repairs, and when they do the company must replace them. Since this outlay occurs at infrequent intervals, it is not feasible to provide for it out of current earnings, and a reserve for depreciation must be accumulated against the time when the outlay must be made. Another but more unpredictable source of depreciation is obsolescence. A public utility ought to give its customers up-to-date service, and to do this it must keep pace with progress in the arts. In order to keep pace a company must sometimes discard or rebuild equipment before its technical usefulness has expired. Examples of equipment that is still technically useful but obsolescent are sleeping cars without air-conditioning and the telephone instrument divided into a receiving and a transmitting piece.

In the matter of depreciation, commissions have encountered two exactly opposite policies in the practices of the utility companies. A majority of the companies doubtless make as accurate guesses as they can, but some companies, the managers of which wish to make earnings appear greater than they actually are, make inadequate annual charges for depreciation; others exceed reasonable expectations of annual depreciation, build up enormous reserves collected from customers, and reinvest them in the expansion of plant, upon which a fair return is again demanded. If depreciation could be predicted accurately, there might be no difficulty about this expense. But it cannot be. The result is expensive investigations, court battles, and discouragement for competent commissioners and the public. Finally, it must be noted that any expense which a company may incur in making a defense against the ruling of a commission is an operating expense and may be recovered from the consumers of its services.

MUNICIPAL-UTILITY PRICES

Since there are five principal municipal utilities,—electric light and power, the telephone, gas, water, and street transportation,—five somewhat different problems of price-making are involved. The most highly

developed is the rate structure of the electric light and power industry, which we shall consider for purposes of illustration.

Part of the business of electrical companies is competitive. Industrial users always have several alternative sources of power. They may install various types of prime mover,—steam engine, Diesel, gasoline motor and sometimes waterwheels,—or they may purchase current from a public utility. The power they consume is called wholesale power. If the central electric stations are to get this business they must furnish power cheaper or more efficiently than the industrialist can provide it by owning and operating one of the prime movers mentioned above. Often, especially in industries where large amounts of heat are used, the competition between energy purchased from an electrical company and energy generated by a steam engine is close, and the current will be purchased only if special concessions are made. The price at which current is sold for industrial use therefore is settled in many instances as is any other competitive contract between a seller and a buyer. There are, of course, power rates that apply generally to small users who are not able to install their own prime movers, but large users usually will compel the electrical company to furnish power at least as cheaply as they can generate it in their own plants. On the other hand, the utility will try to get the contract at a rate as near as possible to the maximum the user will pay.

Commissions do not often interfere with the rates charged large industrial users though they have the legal authority to do so in most of the states. Interference usually occurs when these rates are shown to be excessive or when they are shown to be so low that losses are incurred which the company then attempts to recover by charging small users rates in excess of the cost of serving them.

For domestic lighting and power, for most commercial lighting, and for lighting and power used by small factories and shops, competition is either nonexistent (e.g. domestic lighting), or a wide margin of difference in cost and efficiency exists between self-supply and purchased current. It is in these fields that exploitation of the consumer is most likely to occur and therefore that regulation should play its most prominent role. If the utility companies were left to set their rates without any fear of public interference, they would attempt to charge these consumers discriminatory monopoly prices. Not only would the domestic and other small consumers be compelled to pay prices more or less in excess of the cost of serving them, but they would also be charged dis-

criminary monopoly prices.¹ The most urgent uses, such as lighting and other small consumption in the household and the office, store, and small shop, would be charged for at higher prices than the less urgent, such as cooking, water-heating, and the larger motors. When the average price for the domestic consumption of electricity in both England and the United States was about 6 cents a kilowatt-hour, a few years ago, there were isolated instances in both countries of charges of more than 20 cents. Of course, when such rates were imposed, the consumption was limited to the most urgent uses. These facts merely show that very high prices can be charged for urgent uses. But without regulation the companies would not be prevented from charging higher prices for the most urgent uses and lower prices for the less urgent. Regulation has as its objective not only the prevention of excessive earnings by utilities (average rates in excess of competitive levels), but also the prevention of unwarranted discrimination in prices between different uses.

In general, the principle applied by commissions in setting prices is that each use should pay the price (rate) necessary to cover the cost of providing current for that use. The same principle is applicable also to the regulation of the prices of gas, telephonic communication, and transportation by bus and streetcar. For practical reasons, however, this principle cannot be rigidly applied.

Electric current is one of the most accurately measurable forms of energy. Throughout the United States and most of the rest of the world the unit for measurement is the watt-hour or kilowatt-hour.² Accurate meters for measuring the wattage consumed have long been in use.

The simplest method of charging for current is the straight-line meter rate. By this rate each consumer pays the same rate for a kilowatt-hour that every other consumer pays—say 6 cents. Hence a consumer who used 50 kw-hr a month would pay \$3 and one who used 100 kw-hr would pay \$6. Current would be sold as are small articles in a drugstore or at a grocer's. Behind this rate lies the assumption that a kilowatt-hour costs the seller the same amount no matter how many units are sold to a given consumer, when they are sold, and what the variation in consumption may be. These assumptions are not correct for a number of reasons. Consequently, we find, the straight-line meter rate is seldom employed.

¹ See Chapter XII.

² For the definition of the watt, which is different for direct and alternating current, see any elementary textbook on electricity. A current of 1000 watts flowing for the duration of one hour furnishes 1 kilowatt-hour of energy, the common unit of consumption.

In the first place, each consumer causes the company about the same amount of expense for meter-reading, billing, maintenance of service, and distribution system (pole lines, transformers, service connections, and meter). Hence there is a basic commercial and distribution charge that every consumer should pay. In the second place, each consumer occasions a fixed capital investment by the company. Roughly that outlay is measured by the maximum amount of current that a consumer takes at any moment during the year. That is, if at any moment he turns on enough appliances so that the current passing through his meter amounts to half a kilowatt, he has occasioned an outlay for generating equipment that will produce 500 watts. If the cost of generating equipment is \$200 a kilowatt, then the consumer should pay the company annually, say 10 per cent on \$100, or \$10, because of the fixed charge or *demand charge*. Actually he would not be required to pay so much as that, because the maximum demands of the thousands of domestic customers of a large company are unlikely to occur at the same time. Owing to this diversity, the charge may be reduced measurably. In the third place, the consumer should pay the company for the direct, or prime, cost of generating the current. This is called the *energy charge*.

From the foregoing statements we conclude that a consumer who has a maximum demand (diversity considered) of half a kilowatt and makes that demand at or near the time when a majority of other users make their largest demands, but who uses little current during a month should pay a higher price for each unit of energy than a consumer the peak of whose demand is not coincident with the peak demand of a majority of other consumers and who consumes a great deal of current.

In the following hypothetical table the two cases have been illustrated. Consumer A has the same demand as Consumer B, but because he has a greater variety of use and uses current more continuously his consumption is very large, say 200 kw-hr a month. Consumer B makes his maximum demand at the time when the central station has to carry its greatest load and he uses relatively little current, say 50 kw-hr a month.

TABLE 20

COST ITEMS (MONTHLY CHARGE)	CONSUMER A	CONSUMER B
Commercial and distribution costs	\$1.50	\$1.50
Demand charge ($\$10 \div 12$)83	.83
Energy charge @ 2¢ per kw-hr	4.00	1.00
Total cost	\$6.33	\$3.33
Cost per kw-hr of consumption	3.16¢	6.66¢

As a matter of fact, neither commissions nor utilities have cared to set prices exactly in accordance with cost. The rate structures are so arranged that some of the commercial and demand charges are loaded onto the energy charge. That is, they are made variable, and consumer A will pay more than 3.16 cents and consumer B less than 6.66 cents. The reason for this breach in equity and logic is the fact that a rate structure that compelled small domestic consumers to pay a great deal more per unit than large domestic consumers would be exceedingly unpopular and might provoke political reprisals. It has even been asserted by some authorities that where a single system includes both urban and rural consumers the former should pay a part of the cost of the service to the latter. Several arguments are made for this policy. One is particularly significant. It is said that the use of electricity is a great civilizing influence somewhat similar to education and that the cost of providing it should be borne according to ability to pay and not according to cost of service. This argument, which has not as yet gained wide acceptance in the United States, really rests on the tacit assumption that when the government steps in to regulate prices it should not apply the same criteria that prevail in the world of competition. In Ontario the government has subsidized the building of rural electric lines, which means that the taxpayers have contributed to the cost of serving rural communities. This is probably a more equitable procedure than that of loading the cost of the subsidy on the city consumers, some of whom may have lower incomes than the rural purchasers of current. Public services furnished at prices below cost are usually supported from tax revenue and not from unequal charges against consumers.

RAILROAD RATES

The same principles of regulation apply to railroad rates as apply to municipal utilities. The conditions with which regulation is confronted are much more complex, however, in the case of railroads. In the first place, railroad mileage was very great and rate practices were well established when the Interstate Commerce Commission was created in 1887. Before that date state regulation had been applied sporadically, but it had not accomplished much in the direction of imposing on the railroads, or in inducing them to accept, a rational rate structure. It was not until many years after its inauguration that the commission received enough power to make its decisions effective. Hence the railroad rate structures had become more or less fixed before public interference was imposed.

These structures are not only complicated in the extreme but they are also intimately woven into the industrial structure of the country. For example, a road that had insufficient traffic in its early history sometimes made special rates to induce an industry to locate at some point on its line. Without such concession the industry might have located elsewhere. Once established, such an industry often depends on the favorable rates given at the time of its inception. A regulatory commission cannot undertake the rearrangement of industry within the country. Hence rate structures are only partially modifiable by it.

In the second place, the railroads were and are much more competitive than municipal utilities. If two of the latter compete they are nearly certain to compete over the same field, whether the entire city or a section of it. But two railroads may compete between certain points and still have much traffic over which they have a large degree of monopoly control. The existence of this noncompetitive traffic enables them to continue to compete while two electric concerns would either ruin each other or combine.

In the third place, the types of service rendered by a railroad are much more numerous and much more complex than are the services rendered by a municipal utility. A railroad carries thousands of kinds of goods as freight in addition to mail, express, and passengers. It carries these goods for many types of shipper, in lots ranging from carloads to small packages. It must perform auxiliary services, such as refrigeration, heating, switching, and loading and unloading small shipments, in connection with the transport of commodities.

In the fourth place, the transport of goods is much more a service for industry than are many if not most of the services of municipal utilities, which receive the major part of their revenues from commercial enterprises, municipalities, and domestic consumers.

In general, the same principles govern the regulation of the earnings of railroads as govern the regulation of the earnings of municipal utilities. The railroads are entitled to a fair return on the fair value of the property used for the benefit of the public. But the determination of specific rates for each of the thousands of commodities carried is much more difficult. It is, in fact, often impossible to allocate even the operating expenses to particular shipments. The phenomenon of joint cost appears in relation to traffic moving in opposite directions over the same line, and, since a certain number of trains must be run in order to maintain the service and attract customers, the allocation of expense to different items moving in the same direction is mostly by the process of estimate.

The commission has generally allowed the customs developed by the railroads to stand. In the Eastern territory, goods in the same classification pay rates in accordance with the distance shipped, though with less than proportional additions as distance increases. But this principle is not strictly applied in Southern territory. On many commodities shipped between the Pacific coast and points east of the Rocky Mountains the rates are somewhat similar to postage rates on letters, owing to the competition of the water-borne freight passing through the Panama Canal, which is, of course, carried much cheaper than by land.

Commodities that have high value in small bulk usually pay a higher rate per ton-mile than do bulky goods. The principle applied is not that of cost of service but the one made use of by monopolies—what the traffic will bear. The existence of so many rates that do not follow the cost principle does not prove, however, that freight rates are discriminatory monopoly prices. When a lumberman logs off a territory on which are trees of different species and of different sizes he must cover all expenses. He does this by refusing to sell his products until the prices received for all of them are sufficient to cover costs. But he need not allocate, indeed he cannot allocate, all items of cost to the different kinds of lumber or to the different grades of one kind. To illustrate simply: the wages of the man who fells the trees, the wages of the sawyer, and the capital costs of the sawmill cannot be allocated on any rational basis to the several grades of timber into which the tree can be sawed. The relative prices charged for the different grades of lumber depend on the relative demands for them. So too with the railroads, the principle of cost is to a large extent inapplicable. Hence in the course of the development of the railroads, which often were built ahead of the demand for transport, the goods that would move at high rates were charged high rates and then special concessions were made to develop additional traffic. The industries that located along the routes of the railroads making these special concessions were often unable to compete in the markets without these special rates. Public opinion and political pressure joined with the financial interest of the railroad in maintaining these rates. Certain bulky freight, such as grain, coal, and lumber, was prized by the roads because of the volume it provided and was therefore given concessions.

In 1913 Congress passed an act requiring the Interstate Commerce Commission to make and keep up-to-date a valuation of all the principal railroads of the country. In 1920 another act directed the commission to establish such rates as would enable the companies to earn a fair return on the *aggregate* railroad property of the country as a whole or on the

aggregate within such rate groups or territories as the commission should designate. The commission, however, was permitted to exercise some discretion in respect to particular rates.

Since some railroads in the same territories have more traffic than others, have shorter or less expensive lines between important terminals than others, or for other reasons can carry traffic more cheaply, a uniform system of rates that affords the weaker roads a fair return will give the stronger excess profit. Yet for various reasons rates must generally be uniform over relatively large territories. Hence the act of 1920 provided that all earnings of any company over 6 per cent on the value of the property should be divided into two parts. One half of this excess was to be paid to the commission and placed in a contingent fund from which loans were to be made to weaker roads. The other half was to be retained by the road that earned it but held in reserve, and from it the road might make payments for interest and dividends in any year that earnings fell short of 6 per cent.

These provisions of the act of 1920 were found impracticable for the following reasons: (1) prolonged litigation developed, as might have been expected, with reference to the value of the properties; (2) motor carriers and private automobiles cut into the business of the roads; (3) there was constant pressure from the farming districts, chiefly in the Middle West, for lower rates on agricultural products.

The depression following 1929 caused a great decline in the traffic and revenues of the roads. Some were faced with receiverships, and all large systems failed to earn a fair return. The pressure for lower freight rates for the relief of depressed economic groups continued. The Federal government came to the rescue of the distressed railways with loans, and with special legislation for arranging the liabilities of creditors.

In 1933 the Emergency Transportation Act was passed. This statute is comprehensive. Here we are concerned, however, only with the parts dealing with rates. Certain provisions of the act of 1920, including the clause requiring the recapture of earnings, were repealed. The rule for rate-making was changed. Instead of the relatively inflexible formula of "a fair return upon the value of the property," a more flexible procedure was set up. The commission was directed to give consideration to the effect of rates on the movement of traffic and to the need of the public for adequate and efficient service at the lowest rates consistent with the cost of the service.

This series of statutes and the administrative acts pertaining to them and other statutes makes the Federal government virtually, though not

technically, a partner in railroad transportation. The commission is largely responsible for the rate structure; the National Railroad Adjustment Board selected by employees and carriers has limited jurisdiction over labor disputes; and the government has developed a policy of coming to the relief of roads that are in danger of becoming bankrupt. It is evident also that for the present and near future freight rates are to be determined in part by the effect of these rates on the economic welfare of economic groups and territories. The cost-of-service principle toward which rate-making appeared to be tending in 1920 has been pushed farther into the background. It has not been formally abandoned, and the companies are still primarily responsible for the management of the service and financing, but the government has assumed the major responsibility for rates and has intervened in financial and labor matters.

SUMMARY AND CONCLUSIONS

The attempt of the Federal and state governments to regulate the prices charged by municipal public utilities and railroads has been an attempt to use competitive standards for the regulation of prices that otherwise would be monopoly prices. This policy has not been wholly successful. The criterion for determining competitive standards—cost-to-serve—was found difficult to ascertain and, therefore, not completely applicable. It remains the chief criterion for the regulation of municipal utility rates, but there are many modifications arising out of the difficulties of allocating joint cost and out of questions of public policy.

The regulation of the prices of the services of railroads is now tending away from the competitive or cost-to-serve standard. For the present at least, broader standards of public welfare have overborne that standard.

AGRICULTURAL PRICES

Until recent years governmental interference with prices and production was confined chiefly to the regulation of the charges of monopolies, which was the subject of Chapter XII and the preceding pages of this chapter. There was, however, one very important exception to this general practice. By means of the protective tariff the prices of protected commodities were raised, and the types of production within the country were changed. Moreover, it was the purpose of the tariff to bring about such results. When, after the close of the World War, the agricultural bloc became politically powerful, the precedent created by tariff legisla-

tion in favor of industrial interests was pointed to as justifying government interference for the benefit of the agricultural classes. Of course the tariff had been applied to agricultural commodities, but because so large a part of them had been sold in foreign countries before 1930 these import duties had been ineffectual in respect to most farm products. Hence there was a demand for some type of legislation that would "make the tariff effective for agriculture."

What the agricultural class advocated was a policy that would raise the prices of its major products relative to the expenditures of the farmers, so that they might have larger real incomes. Merely raising the general price level could not bring this about, although it did reduce the burden of interest on indebtedness. A general rise in prices could not have altered the relation of the prices of farm products to the prices of the products of manufacture and could not have raised sufficiently the real incomes of this class. It was evident that only a greater rise in agricultural prices than of prices in general would satisfy the demand. Obviously, this result could be brought about in several ways. If the prices of manufactured goods were reduced while the prices of farm products remained stationary, that would bring the required result. This might come to pass through the reduction of the costs of manufacture. But it was not feasible to reduce the prices of manufactured commodities. In fact these prices were already too low in the opinion of those in authority.

The attempt was made to increase the demand for the farmer's products by measures designed to revive industrial activity in the cities. These measures included the President's Re-employment Agreement, the activities of the National Recovery Administration, the loans of the Reconstruction Finance Corporation, and the legislation affecting banking and the currency.

The low prices for farm products were attacked directly in a number of ways. The surplus of swine was removed by the purchase and destruction of several million pigs; some wheat was exported under bounty; and a quantity of cotton was taken off the market by allowing the owners to store it and borrow money from the government against the warehouse receipts. The principal attack was by means of restriction on production. Measures presently to be described were adopted to induce or to force the growers of wheat, corn, cotton, tobacco, and some less important crops to reduce the acreage devoted to production.

In the case of cotton and tobacco, prohibitive taxes were levied on the processing or sale by any grower in excess of certain predetermined

quotas. But the main reliance was placed on voluntary co-operation. The government offered to pay each grower who agreed to curtail the acreage seeded by a definite percentage a benefit for each acre diverted to other uses. The growers who wished to participate in these benefits entered into contracts to curtail plantings. A large majority of the growers of all these crops entered into the contracts calling for restriction. As a result of these restrictions, the devaluation of the dollar, reviving urban business, and the drought in 1934, the prices of the products were raised relatively to manufactured commodities. In 1936 the Federal Supreme Court declared the AAA restrictions unconstitutional. Then the government introduced a new measure, which has not as yet been passed upon by that court. By this plan farmers agree to restrict production of certain crops and to plant the acreage thus released to other crops that conserve or increase the fertility of the soil or prevent erosion. The government still pays a rental benefit for each acre withdrawn, but the source of the funds thus employed is the general fund of the treasury instead of the processing taxes imposed under the AAA.

Under the AAA arrangements a processing tax was levied on the manufacturers who utilized the four main crops, except that the tax was levied not only on the processing of corn but also on the processing of hogs by the packers. Obviously, a tax only on the processing of corn would have yielded insufficient revenue because so large a part is fed to livestock. The proceeds of these taxes then were used to pay benefits to farmers who co-operated in the restriction program.

In general it may be said that in consequence of the restrictions, prices were raised to the ultimate consumers. This of necessity resulted in sales resistance varying in degree according to the different elasticities of demand for different products. In so far as prices were raised, the volume of consumption was of necessity less than would have been the case had there been no benefits and no processing taxes. By and large, it is the view of competent agricultural economists that reduced consumption at higher consumer prices followed from the act.¹

The actual reduction of crop yields, owing partly to the droughts in 1934 and in 1936, raised prices, and the consumers bore the burden of these increases. To this generalization there was one exception. The domestic tobacco crop is purchased in the main by a few large processors who are in a monopolistic position. Moreover, much tobacco is held in

¹ Joseph S. Davis, *Wheat and the AAA* (1935), p. 364. Also Edwin G. Nourse, Joseph S. Davis, and John D. Black, *Three Years of the Agricultural Adjustment Administration* (1937), pp. 408-411.

storage for more than a year, and the accumulated stocks of the large manufacturers of cigarettes are always great. Farmers' receipts from tobacco fell from 174 million dollars in 1929 to 68 million in 1932. The profits of manufacturers rose, however, from 134 million dollars to 146 million.¹ The manufacturers did not reduce prices to the consumers by the full amount but, like monopolists, took the difference in profits.

It has been said that the prices of the finished products were largely independent of the prices of the raw materials. Like all goods for which the material costs vary more frequently than the prices of the products and often in the opposite direction, tobacco products offered an excellent opportunity for "administered prices." Hence the increase in the cost of tobacco was not reflected in the price of the finished products. The return to the growers increased from 68 million dollars in 1932 to 128 million in 1933 and 159 million in 1934, while the profits of the manufacturers decreased from 146 million dollars in 1932 to 80 million in 1933, and then rose to 100 million in 1934.²

The price of wheat also increased markedly from 1932 to 1934. But many factors combined to cause this increase from 52 cents a bushel (Chicago base price) to 95 cents. Monetary influences, the drought, and the return of a larger measure of prosperity were the major influences. One writer who has made an exhaustive study of the problem concludes that the restrictions of the AAA probably did not hold the price of wheat more than 2 or 3 cents above what it otherwise would have been in 1933-1934. In the following year the effect was problematical.³ Similarly for cotton and hogs, the results were not startling.

It must be remembered, however, that the AAA did succeed in reducing acreage, and that 2 or 3 cents a bushel for wheat and 1 or 2 cents a pound for cotton amount to large sums on the total crop of the United States. Moreover, if the prices of these commodities could be raised by that amount, it is easily possible that they might be raised still more by further restriction.

We have seen that in regulating public utilities and railroads the government encountered very great difficulties of administration. The AAA was in operation too short a time for much to be learned about the obstacles that would be encountered if the government entered into a permanent policy of regulation. Some of these difficulties had already made their appearance in 1934, and others might well have

¹ Harold B. Rowe, *Tobacco under the AAA* (1935), p. 84.

² Rowe, *op. cit.* p. 222.

³ J. S. Davis, *op. cit.* pp. 361, 364.

developed as time went on. Voluntary co-operation was not a complete success in the wheat program. In some states where wheat was a minor crop there was very little acceptance of the plan. There was dissatisfaction with the methods by which the quotas were assigned to different farmers. The wheat area has never been stable, and neither have the corn and cotton areas. No plan has as yet been devised for both control and an allowance for change of area. In the case of regulation of public utilities the courts finally worked out a criterion or standard for regulation which has the support of economic theory, even though its administration is difficult. One may ask, What criterion was set up for determining the amount by which the prices of farm products should be increased? At first the farm leaders asserted that the prices of these products should be raised to the same level relative to all commodities that had obtained on the average from 1910 to 1914. But no very convincing argument has ever been advanced to support this standard, except possibly one that will hardly be accepted as a general rule for all regulation, viz.—that at that time the relative prices of farm products were very favorable to the farmers. It would not be an acceptable argument if used to defend public utility or railroad rates. In the Soil Conservation Act another standard was set up. By this act the farm population is supposed to receive the same real income relative to the nonfarm population that it received in the prewar period. But in the prewar period there were no rayon, no radio, few hard roads, few rural telephones, little central heating in farmhouses, and few consolidated rural high schools, and the automobile was a very different machine from the car of today. To do more than make an intelligent guess in determining what prices would give the farm population the same relative real income is impossible.

Other forces than statistical measures of relative real income are likely to play a more important role in determining our agricultural policy, among which are the number of people who decide to remain on the farms, the progress of farming technique, certain broad concepts of social necessities and political strategy. In these matters regulation is even farther from an accurate standard than it is in the case of public utilities and railroads, though that does not mean that we shall see an abandonment of attempts to control the future of agriculture and the relative incomes of those who live by it. The advocates of protection have tried for many years without any measurable success to find a basis for "scientific tariff rates," and yet protectionism has been the policy of the United States for more than three quarters of a century.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Are the following natural monopolies: the supply of manufactured ice? of natural gas? of milk for urban consumption? of taxicab service? Are they regulated?

2. Suppose that the freight rates on agricultural products and on valuable manufactured goods were now made the same, which would mean higher rates for the former. What change in the location of livestock production in the United States would probably take place?

3. There are at present five evening passenger trains operated by five companies between two important cities. Only in very busy seasons—a small fraction of the time—do these trains carry as many cars as the engines and crews could handle conveniently. Why, do you suppose, the excess capacity continues?

4. Suppose that all land in production used in the United States were to be appraised at reproduction cost. What would the resulting valuation be—higher or lower than the present sale value of the lands? Suppose that the houses in a typical residential district also were appraised at original cost, would the result exceed or be less than sale value?

5. As the law now stands, the prices charged by a utility may not be set so low that the investors do not receive a fair return on the fair value of the property. The prices charged by grocers are regulated by competition. Do you believe that this force always allows a fair return on fair value? more? less?

6. It has been said that to restore farm prices to the same level relative to all prices as existed from 1910 to 1914 neglects changes in the cost of producing both agricultural and manufactured products. Why?

CHAPTER XVII · Speculation and Prices



PRODUCTION FOR THE FUTURE

The production and sale of goods under modern industrial conditions and for modern markets, which sometimes embrace the entire world, always involve more or less uncertainty. Cotton-manufacturers who produce for the domestic or the foreign market cannot foresee exactly the price at which they will sell their cloth; the merchant who lays in a stock of goods cannot tell in advance the prices at which his patrons will buy; and the farmer cannot know in advance the demand that will exist for his products when they are finally brought to the market place. Inability to forecast demand is thus seen to be one of the prime causes of business uncertainty. And this uncertainty is partly due to the widening of the market. When production was carried on to supply a local, restricted demand, uncertainty was far less. Much higher powers of business judgment are necessary now than when most enterprises drew their raw materials from sources near at hand and sold their products to consumers living in a relatively small geographical area. Another cause of uncertainty is the fact that capitalistic production, involving, as it does, the use of much fixed, durable capital, necessitates planning ahead for relatively long periods of time.

Not only is the demand for goods uncertain, but the total market supply likewise is not definitely predictable. In agriculture and, to a less extent, in other extractive industries the annual product is dependent upon natural conditions, such as the weather, which are beyond man's control. Nor is the capricious behavior of nature the only source of uncertainty in production. Workers may strike without warning; credit may be expanded or contracted, owing to a change in the policy of the banks; business may become dull under conditions that at another time seem not to affect in the least the upward trend of output, simply because at one time a majority of people are in a cautious mood, whereas at another time they are sanguine.

All the types of uncertainty that have been enumerated are purely "business risks." But these do not make up the entire list of uncer-

tainties surrounding the conduct of a large enterprise. There are risks due to fire, storm, and other natural hazards, which impinge on the community generally, including the business enterprise. But although these hazards affect prices, they are to a great extent insurable and can be transferred by the businessman to insurance companies. Through insurance the cost of replacing destroyed property is spread over the industry or community, each business or each locality paying its proportionate share for the sake of escaping a less certain but very much greater possible loss.

METHODS OF MEETING BUSINESS RISKS

If future events could be predicted accurately, entrepreneurs would be able to calculate the future costs of making good their losses, and these losses would be added to costs of production or subtracted from gross income. Prices would be affected by a change in the magnitude of the losses, to be sure, but there would be no business risks. As things are, it is not the total amount of the losses that embarrasses the businessman but the impossibility of being prepared to meet them.

One obvious method of eliminating uncertainty, and therefore risk, is the perfection of knowledge of future events. Just to the extent that the cotton-manufacturer knows in advance that the foreign demand for his products will be small, he avoids loss from overproduction for that market. And if the merchant ascertains in advance that sales in his territory are certain to be reduced during the coming season because of dullness in other lines of business, he can escape the loss of having to sell a part of his stock at less than cost. All methods of increasing the amount of information available for the use of business in forecasting future demand and future supply must tend to reduce risks. In the United States there are four types of service that tend to reduce the uncertainty of business: (1) The Federal government collects and publishes a great deal of current data concerning the acreage, condition, and yield of crops, the general conditions of business in foreign countries, and the trend of foreign markets.¹ It publishes also information as to the output of factories and mines, prices, and financial data at short intervals, and an annual summary of the same sort of data together with information on power used by producers, the number of workers employed during the year, and the value of the products of important

¹ See *Crops and Markets* and *The Agricultural Situation*.

industries.¹ (2) Certain informational services are maintained by private concerns and sold to subscribers. Some of these services are general in nature; others pertain to particular trades or groups of trades. (3) Trade and financial journals collect and disseminate information, partly in the form of descriptive statements to the effect that sales are "poor" or "good" in different localities or different trades, or that collections are "slow" or "good." Much statistically descriptive information also is included in these statements. (4) Some large concerns maintain their own informational services. Facts are collected, in quantitative form when possible, and analyzed, and conclusions are drawn for the use of the managers.

When any concern can obtain more accurate information about future conditions than its competitors, it can make greater net gains and, as far as competition is complete, drive them out of business. By avoiding losses it can reduce costs of production; and by knowing when and how much to produce, it can better serve its market. Society also gains from the elimination of the losses incident to business uncertainty; for if one competitor can reduce risks by perfecting its information concerning future markets, its competitors will imitate its methods, costs to marginal producers will fall, there will be less waste of the factors of production, and the national income will be increased.

Still another method of reducing uncertainty is by consolidation. A producer who sells his wares over a limited geographical area must, in the absence of foreknowledge of the variations of demand and supply in that market, share all the ups and downs of general business in that area. But some of these variations can be escaped by selling over a larger area. Many of the causes of variations in demand and supply are local in origin and effect. Partial crop failures and overproduction are often limited to a small region. A merchandising concern that has many stores scattered throughout a country, or a packing company that sells meat over many districts differing in resources and industries, often can take advantage of its wide distribution by selling in other districts goods which would normally be demanded in one district but which cannot be sold there because of some unforeseen disturbance of local incomes; or, if competition becomes so severe that costs cannot be covered in one place, the seller can withdraw or temporarily cut down his supply and dispose of his products in places where the reverse conditions prevail.

When many risks are thus consolidated, or brought together and

¹ See the *Survey of Current Business* and the *Commerce Yearbook*.

assumed by a single concern, there is a gain in predictability. The company that insures dwellings against fire can foretell very little better than the owner the chances that a single house will burn down in the course of a year; but it can predict with much greater certainty how many houses out of several hundred thousand scattered over a wide area will be destroyed. It can do this because the chance conditions that cause fires follow a normal course through the years, and a greater prevalence in one locality is offset by a lesser prevalence in another. In the same way, the large business concern that can distribute its sales over a wide area or that can produce several noncompeting commodities can rely on the greater constancy of phenomena when a large number of cases are dealt with.

SPECULATION AND RISK-TAKING

When future events are uncertain, different people are likely to have divergent opinions about them. This is true of future prices. Vacant parcels of land in a growing city are almost certain to be used at some time in the future; but the exact time at which they will be improved and the value they will have at that time are uncertain, and among a group of real-estate dealers who know the local conditions there will almost certainly be some difference in estimates in respect to these two important future events. The present owners of these tracts of land may believe that future values warrant a present price of \$5000 a lot. If other persons believe that this price is too low, they will probably buy the lots and wait for the future rise in values. In doing this they are engaging in speculation. This is the simplest and most ancient type of speculative dealing; buying a good outright in the expectation of making a gain through a rise in price that is greater than the rise expected by the seller. Speculative deals of this sort, of course, are exceedingly common; but usually they do not attract much attention because they are mostly confined to local markets, because their volume is unknown, and because they have the sanction of long-established usage.

The speculation about which the public is concerned today is that which takes place in the stock and produce exchanges. These are large markets; and because they are large the total volume of sales attracts the attention of the entire country. They are not, of course, confined to speculative operations. An investor who wishes either to buy or to sell stocks and bonds finds in the New York Stock Exchange a continuous market for listed securities. The owner of a country elevator

or the farmer may sell his wheat on the Chicago Board of Trade, and the miller who wants wheat to supply his mills may buy in the same market. But by far the larger portion of all sales in these two types of market are speculative deals.

In the produce markets the speculation takes the form of "dealing in futures," which is distinguished from "cash sales." In the latter, grains and other agricultural products are bought from samples, and the goods are actually delivered to the buyer; in the deal in futures the seller merely enters into a contract to deliver, say, wheat of a certain grade to the buyer at some future date. Such a bargain takes place just as it does in the real-estate market—because two persons happen to have different estimates as to the probable future price. The seller believes that May wheat should sell somewhat below \$1.50 a bushel; the buyer believes that it should sell above that price. All those who believe that the price will be less than \$1.50 and want the market to go lower are called bears, and their operations are said to *bear* the market; those who believe the market will go higher and want it to go higher are called bulls, and they are said to *bull* the market. In a typical deal in May futures the buyer enters into a contract with the seller to take from him 10,000 bushels of wheat at \$1.50 a bushel, in May, say. If the seller can go into the cash market and buy wheat of the stipulated grade for \$1.40 a bushel when the time for fulfilling the contract approaches he will make 10 cents a bushel on the entire lot, and the buyer will lose the same amount of money. But if the seller can cover his contract only by buying at \$1.60 a bushel, he will lose, and the buyer will gain.

As a rule, no wheat is delivered on such contracts. Instead, the seller of futures will buy up a contract to buy and turn it in to liquidate his obligation. Both the buyer and the seller may change their positions several times before May arrives, buying at one time and selling at another. The net effect of such transactions where actual delivery is not even contemplated, is simply that whoever finds that his contracts, if carried out, would bring him a loss must pay that amount to others whose contracts would entitle them to a gain.

On the stock exchange speculative transactions have the same general features, with the difference that future contracts are not entered into, being prohibited by the rules of the exchange. Instead, the seller goes to the loan counter and "borrows" the stock in order to make actual physical delivery. He must deposit a sum sufficient to be security for the borrowed stock at its market price. The deals take place for the

same reasons as on the grain exchanges. The seller believes that a given stock is selling at a price that the future prospects of the issuing company or general business conditions do not warrant. He believes that the stock will go down in price; the buyer believes that it will rise. If the stock goes down, then the seller can "cover" (that is, replace the shares borrowed) at a lower figure, collect the difference between the price at the time he sold and the then ruling price, and pocket the difference; but the buyer loses. If the stock rises in price, then the buyer gains and the seller loses.

Speculative transactions in both grain and stocks exhibit some resemblance to gambling. But there are significant differences. The chance on which the gambler bets is created by him for the purpose of gambling. The chances, or uncertainties, on which the traders in the exchanges risk their money are not of their making. As we have seen, they are the unavoidable concomitants of modern business methods. They persist because men are unable to forecast changes in demand and supply. In the second place, the gambler does not even incidentally contribute to the production of goods, whereas the speculator may do so.

EFFECT OF SPECULATION ON PRICES

When a speculator in lands buys in the expectation of a greater rise in value than is commonly anticipated in the market, he raises the demand for such land and he also tends to raise the price of similarly situated parcels. If he is inexperienced he may, of course, be entirely mistaken, and his estimate will have little effect upon the market unless he is only one of many such amateur traders. In the latter case a boom may develop which will benefit no one but the original owners of the property. The amateurs will lose money and the community will be poorer only by the amount of time and expense wasted in carrying through the deals. The losses of the buyers resulting from overestimate will be exactly offset by the gains of the sellers. On the other hand, if the speculators are experts it is likely that their guesses as to the future values of the land will be better than those of the owners whose experience and training have been in some other occupation. In that case there may easily be a social gain. For example, the owners of a tract of land may contemplate selling it to a warehouse company. But a speculator who believes that within a few years the land will be more valuable for retail shops outbids the warehouse company and holds the land for the more valuable use. If this forecast of the future develop-

ment of the city proves correct, then both he and the community will have gained as a result of the speculation.

In a speculation in wheat a buyer may believe that the November price for wheat deliverable in May is too low. His knowledge of the market leads him to believe that the crop forecasts in the Argentine are too optimistic or that the demand will be heavier than has been supposed. He concludes that by May the price will be higher than it now is on future contracts, and so he buys for delivery at that time. If his guess proves correct, he will make a profit. But has his action been of any benefit to the community? In the first place, the fact that he has bought has increased the demand for May delivery; and if his purchases are large enough, the price will be forced up on that account. And the rise in the May price will affect the cash price in November. The result will be a more equal distribution of the wheat for use during the year and a more nearly uniform price throughout the year. To be sure, these obviously desirable results are realized only if the speculators tend to make more correct estimates than other people. But the bulk of the trading on any large grain exchange is done by experts, and the estimates of future demand and supply which they make are much more likely to be correct than are the guesses of outsiders or even of farmers and local grain-buyers. To this extent the existence of the organized speculation results in social gain.

There is only one way to avoid the risk of price changes, and that is by consuming as soon as production has been completed. The farmer who harvests a thousand bushels of wheat must sell at once if he is to avoid risk. If he holds his grain, it may either rise or fall in price. But not all farmers can sell their crops as soon as harvested. Even if they could, the crops would not be immediately consumed, and the dealers who bought and held them would be compelled to assume the risk. It is the primary business of a farmer to grow grain and of the local buyer to gather it together into carloads, clean it, and dispatch it to a central market. Even the warehouse company that stores it at the central market has a special occupation, and so has the miller who grinds it into flour. But the risk of price changes exists, and in the absence of a specialized body of risk-takers, it falls on the person who holds the grain. It is the function of the traders on the exchange to assume most of the risks of price changes. Since they are experts, it may be supposed that they will make fewer mistakes than any of the other persons concerned in the production and distribution of grain and its products.

The risk of variations in price can be transferred to the specialists in

risk-taking by the simple device of a hedge. A farmer whose wheat is in the granary by September may be unable to haul it to market until three months later. If he wishes to avoid the risk of changing prices, he can sell December wheat in the future market and leave his wheat in the granary until it is convenient to deliver it to the local elevator. If the price falls lower than he had anticipated, he will be able to cover the loss on his crop by the gain on his transaction in the futures. The manager of a local elevator may cover the risk of price change while his shipments are in transit by selling futures in the exchange. The miller who has filled his warehouses with grain stands the risk of losing money if the price of wheat (and consequently of flour) falls before he has milled and sold the wheat. He too may escape risk, transfer it to the traders, by selling wheat for future delivery.

It must not be inferred that all the dealing on the grain exchanges contributes to desirable ends. When risks are assumed by those most competent to take them, and when the dealing in the exchanges causes prices to fluctuate less than would otherwise be the case, there is a social gain. The more even distribution of consumption in accordance with need and in spite of periodic supplies also is a social gain, for if too great consumption at abnormally low prices occurs at one time, it must be offset by diminished consumption at higher prices at a later time. We know from our study of the principle of diminishing utility that the satisfaction consumers derive from an additional block of goods does not offset the deprivation in utility suffered from a shortage later on. As the amount consumed is diminished, the total satisfaction enjoyed by the users diminishes more than proportionally; and as the amount increases, the total satisfaction increases less than proportionally.

THE REGULATION OF THE EXCHANGES

However beneficial may be the effects of speculation when it is carried on under ideal conditions, it is clear that it may be perverted from socially beneficial uses. Both the produce exchanges and the stock exchange are often manipulated by traders for their own selfish purposes and to the detriment of the public. Because of this fact both have been subjected to regulation in recent years. The first step in the direction of regulating the produce exchanges came in 1922 when Congress passed the Grain Futures Act. This law imposed virtually no regulation, but it did require the reporting of many transactions to the Secretary of Agriculture. The agitation for regulation grew. It was based upon two complaints.

In the first place, it was commonly believed among farmers that the large operators in futures were able to depress or raise the prices at times when it was to their advantage to do so and that these operations were harmful to the growers of grain. In the second place, it was believed that the markets were sometimes manipulated by unfair methods, such as the circulation of false reports, corners, and wash sales.

In 1936 Congress passed the Commodity Exchange Act as an amendment to the earlier act mentioned above. Produce exchanges were brought under the regulation of a Commodity Exchange Commission consisting of the Secretaries of Agriculture and Commerce and the Attorney General of the United States. This body may limit the amount of speculative trading that may be done by any one person on a given day, and it may limit the amount of his holdings—either long or short—at any time. The Secretary of Agriculture was given wide discretionary powers over a great variety of practices.

Wash sales, deception of other traders, circulation of false statements, and the offsetting of orders to buy against orders to sell (no sale or purchase having been made by the broker) are prohibited. All these and similar provisions are designed to limit purely gambling transactions in connection with the great grain exchanges.

The act covers a large variety of farm products, ranging from such important crops as wheat and corn to grain sorghums. Here the evident intent is to protect the growers of all farm commodities from the supposedly evil effects of the manipulators in the great exchanges.

Speculation in the markets for securities stands on a somewhat different footing from that of the grain exchanges, so far as social effects are concerned. The things dealt in are not commodities but stocks and bonds. Trading without delivery, although apparently prohibited by the rules of the exchange for many years, forms the greater part of all transactions. That such an exchange is a desirable adjunct to the capital market cannot be doubted. It offers the best possible means for facilitating investment and disinvestment that has been devised. Yet the transactions that take place in these markets are often distinctly inimical to the economic welfare of the country as a whole. One evil is the effect of purely tactical operations of traders—pools, squeezes, etc.—upon the economic conditions within the country. During the boom period that preceded the collapse of 1929, millions of dollars of worthless securities were issued and foisted upon the public through the exchange and many more millions were lost by investors who were deceived by manipulations within the exchange.

There were probably few cases in which the promoters of these flotations were actually guilty of legal fraud. But in many instances the facts which the investor should have had before him when he made the purchases were either negligently or purposely kept from circulation. In many cases half-truths were contained in the prospectuses. The result was a loss of large sums by many small investors and even by large institutions. It is obvious, of course, that short of complete dictatorial control it will never be possible to prevent individuals from making bad investments. All that can be accomplished directly through government control is to prevent promoters and underwriters from making false statements about the companies and governments whose securities they distribute and to compel them and the responsible company officials to reveal the pertinent facts about the finances of the issuer of the stocks and bonds. This the Federal Securities Act of 1933 attempts to do. Furthermore, it was asserted by the President that the act would revive confidence in securities and thus aid in recovery.

The provisions of the law are too lengthy and too technical to permit anything but the most general statement in this discussion. In the main, the law places the responsibility for both the misstatement of fact and the failure to reveal pertinent facts about the securities on all those primarily concerned with their issuance and flotation—the directors of the issuing company, the accountants who prepare the data for the prospectus, the engineers who may sign statements as to the appraised value of property held as assets supporting the securities, and the underwriters who participate in the distribution of them. As the law now stands it is required that all new securities that are to enter into interstate commerce must be registered with the Securities Exchange Commission and that the registration shall be accompanied by a complete and detailed description of the current financial position of the issuer of the securities, the purpose of issuing them, manner in which they are secured, and many other facts. After the new securities have been registered, a "cooling period" between the registration of the securities and their sale to the public is imposed to prevent high-pressure selling and to give the prospective purchaser time to investigate the issue.

In 1934 the Securities Exchange Act was added to the controlling legislation. It further provides administrative machinery for carrying the legislation into effect. Like the Commodity Exchange Commission, the commission dealing with the securities market has very great discretionary powers. These two acts add to the number of administrative boards and commissions, such as the Interstate Commerce Commission,

the Federal Communications Commission, and the state public-utilities commissions, that have been brought into existence to regulate business practices and the prices of goods and services.

It must not be assumed that the legislation for the control of the marketing of securities and for continuous reporting on them will insure investors against unsound investments. In the first place, the acts do not apply to stocks and bonds already in the hands of the public when they went into force. In the second place, the fact that the securities have been properly registered and approved for sale to the public by the commission does not guarantee their soundness in the future. No amount of regulation will prevent business failure or guard the investor against the effects of shifts in demand and of technical change. In the third place, the acts do not apply to all the investments open to the public but only those, to speak generally, that are designed for national distribution. They do not apply, of course, to the direct sale of property.

Finally, there is very little in the legislation that will prevent, though it may restrict somewhat, the development of a speculation craze like that which spread through nearly all classes of the community in 1928-1929. During that period the published statements of corporations in respect to their earnings did not prevent stocks that had no prospect of paying more than two or three dollars a share from rising to heights where as investments they could not have returned the owner on the price he paid more than half the dividend rate. Like most social-reform legislation, the regulation of the securities market depends for the accomplishment of its purposes on the intelligent behavior of the people for whose benefit it was enacted.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The man who does not insure his house against fire is a speculator; the man who insures is not." Are these statements true?

2. The sales of futures on the Chicago Grain Exchange are several times as great as the cash sales. The men who trade in futures deal in "phantom" wheat, and the supply of this phantom wheat must cause the price of real wheat to be lower than it would be otherwise. Is the foregoing statement correct?

3. The variation in the prices of agricultural products not dealt in on the great speculative exchanges is greater in normal times than is the variation in the prices of grains that are bought and sold on these exchanges. Why should this be true? Would the variations in the prices of such commodities as wheat and cotton be greater if the exchanges were abolished?

4. "During the boom in land prices in Iowa in 1919-1920 there was much speculation. Lands were greatly overvalued, and when the decline in prices occurred many farmers were forced into bankruptcy. This was a great social loss." Argue for or against the statement "this was a great social loss." Was the extravagant rise in prices due to speculation or to something else?

5. "Men speculate because prices fluctuate; prices do not fluctuate because men speculate." Is this statement correct? Explain.

6. If a factory-owner buys a new machine, he cannot be sure that his investment will prove profitable. Is his investment a speculation? Does he take a risk?

7. Suppose that the government should forbid all future trading in the commodity exchanges. Who would then bear the risks incident to production and sale of agricultural commodities? Would it be possible for the government to assume them?

8. What is the advantage, for an investor and not a speculator, in owning a security listed on a large securities exchange? Can a company whose securities are thus listed secure capital on more favorable terms by reason of such listing? Why, or why not?

9. Explain the relation of risk-taking to entrepreneurs' differential costs.

BOOK THREE

=

MONEY AND PRICES

CHAPTER XVIII · Money and Credit



THE FUNCTIONS OF MONEY

In the foregoing chapters we have discussed value and price. In modern societies the value of any commodity is measured in terms of money. Value so expressed is called price. In other words, *price* is the quantity of money for which a commodity is exchanged in the market. Money, therefore, is the yardstick by which we measure value. But unlike other units of measurement, such as the yard or the pound, the monetary unit is itself a fluctuating one. The yard is a unit by which we measure space, and it is of constant length. The pound is a unit by which we measure weight, and the pound is of constant weight. The dollar is the unit by which we measure value, but, strangely enough, it is of fluctuating value. Thus the price of a bushel of wheat may double, but the value of wheat may be exactly the same as before. In this case the unit by which we measure value has itself changed in value. Prices therefore must be distinguished sharply from exchange values. The demand for a commodity in terms of money may have increased greatly, and yet the demand in terms of other commodities may not have changed at all.

Money serves not only as the yardstick by which value is measured in modern society; it serves also as a medium of exchange. Goods and services are not exchanged against goods and services; they are exchanged for money or its equivalent. These, then, are two of the important functions which money performs: first, it serves as a measure of value; second, it serves as a medium of exchange. In addition, as will be noted in a later section of this chapter, certain kinds of money perform a special function in the capacity of government and bank reserves.

MONEY AND CIRCULATING MEDIA

“Money” is a broad term; it includes a great many varieties and species. Not all forms of money serve as the measure of value; only standard money performs that function. On the other hand, there is

in all modern countries a great quantity of circulating media which are not, strictly speaking, money. We must therefore distinguish between *standard money*, *money*, and *circulating media*. It would be more accurate, in fact, to say that standard money serves as the measure of value and that money and money substitutes serve as the media of exchange.

Money is any medium of exchange which passes freely between persons *without indorsement*; or, in other words, money is any medium of exchange which is acceptable on its own account, regardless of the financial standing or integrity of the person who offers it in exchange for goods or services. Checks, drafts, money orders, etc. are forms of circulating media which are not acceptable on their own account but which will be accepted readily if indorsed by someone favorably known to the person to whom they are presented.

MONETARY STANDARDS

There are various forms of monetary standards, of which the most important are the following: (1) the gold-specie standard; (2) the gold-bullion standard; (3) the gold-exchange standard; (4) the managed-paper-currency standard.

Under the *gold-specie standard*, the government engages to coin all gold bullion freely and to redeem all forms of money in gold coin on demand. The standard coin contains a definite weight of pure gold, fixed by law. Gold coins are paid out freely by the Government Treasury and by the banks in exchange for other forms of money. Gold coins (or gold certificates, which are in effect warehouse receipts for gold) thus circulate freely in the community. Individuals, if they wish, may legally hoard gold as a store of value. This was the monetary standard in the United States before the suspension of specie payments on March 6, 1933.

The *gold-bullion standard* prevailed in England from May, 1925, to September, 1931. Under this system gold coins did not circulate, and there was no obligation to redeem bank notes and other paper currency in gold coin. Holders of bank and currency notes did, however, have the right to demand gold bullion—bars of not less than 400 ounces in weight—in exchange for paper currency at a stipulated rate. Such gold bars could be hoarded or they could be used in international payments.

With the passage of the Gold Reserve Act of January 30, 1934, the United States adopted a modified gold-bullion standard. Under this

act gold coinage was suspended, and gold coins were not permitted to circulate. Under Treasury regulations, it was possible, however, to obtain gold bullion, in exchange for currency, for the purpose of export or for industrial uses. Under the act the President was empowered to fix the gold content of the dollar within the limits of 50 to 60 per cent of the former weight of the gold dollar. By Presidential proclamation the gold dollar accordingly was reduced in weight from 25.8 grains of gold nine-tenths fine to 15 $\frac{5}{21}$ grains of gold of the same fineness. This established a value of \$35 per ounce of gold. The Treasury stood ready to purchase gold bullion at this rate, and, under Treasury regulations, gold bullion might be obtained for export and for industrial uses at this rate. Such a standard may be called an *international gold-bullion standard*.

Under the *gold-exchange standard*, the central bank is required to redeem its notes on demand either in gold or in foreign exchange of gold-standard countries. In some countries the central bank is required to hold a minimum reserve against its note obligations, a part of which must be in gold and a part of which may be in foreign exchange. As a gold-economy measure, a number of European countries adopted this partial gold-exchange standard in the postwar decade. Some of the more undeveloped countries held no gold reserves at all but simply engaged to exchange the domestic paper currency for foreign exchange of a gold standard country. In this manner the value of the domestic currency was held at a fixed ratio to the currency of gold-standard countries.

The *managed-paper-currency standard* prevails at the moment in England. The pound sterling, the British monetary unit, is not linked in any definite ratio to gold. This does not necessarily mean that the value of the pound sterling fluctuates widely in relation to gold. Indeed from June, 1935, to June, 1936, the fluctuation was less than 1½ per cent. But since there is no definite ratio fixed by law, it is always possible that the gold value of such a currency may change considerably or even violently under the pressure of shifts in the balance of international payments.

KINDS OF MONEY IN THE UNITED STATES

Under the Gold Reserve Act of 1934, a number of important changes were made in the monetary system of the United States. While gold is no longer coined, the monetary unit, the dollar, nevertheless is linked

to gold at the ratio of \$35 per ounce. The government alone can legally hold gold for monetary uses, but the Federal reserve banks can hold gold certificates which are issued by the government to these banks in exchange for gold bullion or gold coin. Imported gold or newly mined gold must be turned over to the United States Treasury, and the gold certificates received in return may be held by the Federal reserve banks. Gold certificates do not circulate in the community but are held by the Federal reserve banks as a reserve against the deposits and note obligations of these banks.

In the following table are shown the different kinds of money in circulation (outside the Treasury and the Federal reserve banks) in the United States as of June, 1936.

TABLE 2 I. Kinds of Money in Circulation

	(IN MILLIONS OF DOLLARS)
A. Treasury currency	
1. Specie	
a. Silver dollars	35
b. Subsidiary silver	316
c. Minor coin	135
2. Paper	
a. Gold certificates ¹	101
b. Silver certificates	955
c. United States notes	278
d. Treasury notes of 1890	1
e. National-bank notes	366
f. Federal reserve bank notes	52
B. Federal reserve notes	<u>4002</u>
Total	6241

This table gives the different kinds of money which are in hand-to-hand circulation and which are held in the tills and vaults of business firms and of commercial and savings banks. The table does not include the money held by the Federal Treasury or by the Federal reserve banks. It will be noted that much the largest items are the Federal reserve notes (about 4 billion dollars), the silver certificates (about 1 billion dollars), and the subsidiary silver and minor coin (about a half billion dollars). The gold certificates, formerly a major item, no longer form a part of our circulating media, the relatively small number (101 million) still outstanding being either lost or illegally hoarded. The silver certificates may be described as virtually equivalent to ware-

¹The small amount of gold certificates still outstanding are either lost or have not yet been returned to the government or the Federal reserve banks.

house receipts for silver. Most people prefer paper money to the clumsy silver dollars. Hence silver circulates mainly in the form of silver certificates, the silver being deposited in the vaults of the United States Treasury as security for the outstanding certificates. The historical and political reasons for the inclusion of silver dollars and silver certificates as a part of our money supply will be discussed in a later section.

United States notes, or "greenbacks," amounting to \$450,000,000 were issued during the Civil War. Some were retired after the war, but the total volume was finally fixed by law at \$346,681,000. During the greenback period (1862-1879) these notes were not redeemable in gold, and were therefore inconvertible paper money. Later they were redeemable and were secured by \$156,000,000 of gold coin and bullion maintained as a reserve in the Treasury to guarantee redemption. After the suspension of specie payments in March, 1933, and later under the Gold Reserve Act of 1934, these notes were no longer convertible into gold.

The treasury notes of 1890, once of considerable importance, have almost disappeared and are now in process of being retired altogether. They were issued under the Sherman Silver Purchase Act, which provided that the Secretary of the Treasury was to purchase 4,500,000 ounces of silver per month. This silver was paid for by the issuance of these treasury notes, which were therefore, in effect, virtually silver certificates.

The national-bank notes formerly constituted a very important part of our money supply. In accordance with Civil War legislation, and until 1935, national banks were authorized to issue bank notes. These were obligations of the banks issuing them and were secured by government bonds of equal value, deposited with the Comptroller of the Currency at Washington, and by a deposit of lawful money equivalent to 5 per cent of the value of notes issued. These notes, moreover, were a first lien on all the assets of the issuing bank. Formerly national banks were required to invest a portion of their capital in government bonds, which fact stimulated the issue of notes. After the establishment of the Federal Reserve System they were permitted to sell these bonds, at the rate of \$25,000,000 per annum, to the Federal reserve banks. On the basis of the bonds so acquired, the Federal reserve banks could then issue Federal reserve bank notes. In March, 1935, the Treasury announced that all bonds bearing the note circulation privilege would be retired on August 1, 1935. The national-bank notes and the Federal reserve bank notes still outstanding are therefore no longer obligations of these banks. They are in process of retirement by the Treasurer of

the United States, who holds a deposit of other lawful money sufficient to redeem the bank notes still outstanding. Those still in circulation therefore now are counted as a part of the treasury currency, that is, currency issued by the Treasury of the United States. When these notes have been retired, treasury currency will consist exclusively (with the exception of a relatively small quantity of United States notes) of silver dollars, silver certificates, and subsidiary coins.

Federal reserve notes are issued by the twelve Federal reserve banks. They are not "lawful money"; in other words, they cannot be counted by the Federal reserve banks as reserve against their outstanding liabilities. How they are secured will be described elsewhere in this chapter.

THE STANDARD OF VALUE

In 1937 we had in the United States about \$12,000,000,000 of gold. Gold has long been, and is still, the standard by which our monetary unit is valued. But gold has not always been our standard money. From 1792, when our first coinage act was passed, to 1834, silver was the actual standard; from 1834 to 1862 gold was the actual standard; from 1862 to 1879 greenbacks were the actual standard; from 1879 to March, 1933, gold was again the standard of value; from March, 1933, to January, 1934, we were again on a paper standard; and since January, 1934, we have returned to a gold standard, though a modified one.

This brief account of our standard money leads us to a consideration of our monetary history. In 1792 we started out with a double standard, both gold and silver being by law standard money. When gold and silver are both made standard money by law, we have what is known as bimetallism. Bimetallism exists when both metals are coined freely; that is, when there is no limit to the quantity of either metal which may be coined into money, any person being free to bring as much gold or silver bullion as he wishes to the mint to have it coined into money. Both metals when coined into money must also be legal tender; that is, both must be accepted by law in payment of a debt unless the contract specifies otherwise. To put it another way, if both gold and silver are legal tender, any debtor may legally discharge his debt by offering to the creditor either gold or silver money. Free (that is, unlimited) coinage of the two metals and equal legal-tender quality are the two conditions necessary for a bimetallic money standard.

In 1792 our silver dollars contained 371.25 grains of pure silver, and all through our history this has been the weight of our standard

silver dollars. In 1792 the gold dollar contained 24.75 grains of pure gold. It will thus be evident that the weight of the silver dollar was fifteen times the weight of the gold dollar. The mint ratio was therefore 15 to 1. This ratio was adopted because on the market one grain of gold bullion sold for approximately fifteen grains of silver bullion. Thus the pure silver contained in the silver dollar was worth approximately the same as the pure gold in the gold dollar. The market ratio was approximately 15 to 1.

But soon the market ratio changed to approximately $15\frac{1}{2}$ to 1. Since the mint ratio was fixed by law at 15 to 1, it is clear that there was not enough silver in the silver dollar to equal in value the gold in the gold dollar. Under these circumstances no one would pay his debts in gold coin. The gold content of the dollar (24.75 grains of gold) would now buy 383.62 grains of silver in the bullion market; and since it required only 371.25 grains to make a silver dollar, one would always be paying the lesser value by paying one's debts in silver dollars. The gold dollar, as a dollar, was worth no more by law than the silver dollar, but the gold content of the gold dollar was worth more than the silver content of the silver dollar. Consequently gold was worth more as bullion than as coin; and as a result it was no longer coined in practice, and coins already made were melted into bullion. Thus the "cheaper" of the two kinds of money, both standard by law, drove the "dearer" out of circulation. This tendency on the part of the cheaper money to displace dearer money is known in economic literature as Gresham's law (Gresham being the name of the man who first clearly stated it). While legally there were two standard forms of money, in actual fact silver became the standard of value.

In 1834 Congress changed the weight of the gold dollar with a view to bringing it back into circulation. The weight of the gold dollar was reduced to 23.22 grains of pure gold; the silver dollar remained as before. This made the mint ratio approximately 16 to 1. The market ratio, however, was slightly lower, let us say $15\frac{1}{2}$ to 1. Under these circumstances no one would pay his debts in silver dollars. The silver content of the silver dollar (371.25 grains) would buy 23.95 grains of gold, whereas it required only 23.22 grains of gold to make a gold dollar. Silver was now more valuable as bullion than as coin. As a result silver was no longer coined into dollars, and existing silver dollars were soon melted down into bullion. Gold became the actual standard of value, and hence we may properly say that the silver content of the silver dollar was worth \$1.03.

In order to keep subsidiary silver in circulation the weight of these coins was reduced in 1853. This reduction accounts for the fact that two half-dollars weigh less than a silver dollar. Being coined in limited quantities, subsidiary silver was unable to drive out gold.

Gold then remained the actual standard of value until 1862, when, under the stress of war, the government issued huge quantities of paper money (greenbacks). These greenbacks were promises on the part of the United States government to pay on demand. But by reason of the excessive issue the promise could not be redeemed; so in relation to gold the greenbacks fell in value. Since greenbacks, though irredeemable, were nevertheless legal tender, everyone naturally preferred to pay his debts with them rather than with gold. The result was that gold dollars disappeared from circulation (except in California) and greenbacks became the actual standard of value.

After 1865 the greenbacks gradually rose in value, and finally (1879) they became redeemable in gold. Thus again gold became the actual standard of value.

In the nineties, however, the gold standard was once again seriously threatened. In 1876 rich silver mines were opened up in Colorado. Silver became cheap compared with gold. The market ratio rose rapidly above the 16 to 1 mint ratio until in 1878 it was 18 to 1. It then became profitable to coin silver into dollars; but the law no longer permitted free coinage of silver, for in 1873 an act had been passed which demonetized silver. This probably had been done without any malice aforethought. Under the operation of Gresham's law, silver dollars had automatically disappeared from circulation as a result of the act of 1834, which established the mint ratio at 16 to 1. When silver became cheaper and the market ratio rose to 17 or 18 to 1, those who had silver bullion (such as the silver-mine operators) wished to be able to convert it into dollars at the mint ratio of 16 to 1. A silver dollar would be as good as a gold dollar in paying off a debt, but the silver content of the silver dollar was now not worth the gold content of the gold dollar. Furthermore, in the period of the seventies the supply of money failed to keep pace with the rapidly expanding volume of trade, prices were falling (as will be explained later), and therefore many people who were suffering from the fall in prices favored a return to bimetallism. Thus arose the silver movement in politics. Congress passed in 1878 the Bland-Allison Act, which required the Treasury Department to purchase silver bullion and coin it into silver dollars of legal tender. But the amount of silver that could be coined under this law in any one

month was limited; in fact, only about 25,000,000 silver dollars were coined each year from 1878 to 1890. These silver dollars were placed in the Treasury, and silver certificates were issued, redeemable in the dollars so deposited.

This act did not satisfy the demands of the growing silver movement. In consequence the Sherman Silver Purchase Act was passed in 1890. This act required the Secretary of the Treasury to purchase 4,500,000 ounces of silver per month, the silver to be paid for by the issue of treasury notes. These were the treasury notes of 1890.

Some 576,000,000 silver dollars were coined from 1878 to 1893 as a result of these two acts. The value of the silver content of these dollars was less than the value of the gold content of a gold dollar; therefore people naturally wanted to exchange these silver dollars for gold dollars if possible. As the treasury notes were by law redeemable at the Treasury in either gold or silver, the silver dollars coined under the Sherman Act were virtually redeemable at the Treasury in gold, since the treasury notes which were based on the silver *were* redeemable in either gold or silver. In practice, therefore, treasury notes were constantly being presented at the Treasury for gold. But although people were anxious to exchange silver dollars for gold dollars, they naturally preferred to pay their debts in silver, the cheaper metal. Gold was thus disappearing from the Treasury on the one hand and from circulation on the other. At this juncture Congress repealed the silver purchase clauses of the Sherman Act and saved the situation for the gold standard. Had the coinage of silver dollars continued, the silver dollar would soon have become the standard of value. The gold standard was established on a firmer legal basis by the Gold Standard Act of 1900, which provided that the Secretary of the Treasury should maintain at par with gold all forms of money issued or coined by the United States.

Thus ended a century of experimentation in the United States with the monetary device of bimetallism. There is only one reason for having a double standard, for having a standard dollar which may be either gold or silver. It is alleged, according to this reasoning, that when two metals are used there is less likelihood that the value of the standard dollar will fluctuate as much as when only one may be coined freely. Now everyone admits that the fluctuation of the standard of value is an evil; and if bimetallism would cure or even mitigate the difficulties arising from this fluctuation, we should all acclaim it superior to monometallism. But the reasoning rests on an assumption that has not stood the test of experience. It assumes that the natural flow of two metals

from the mines will be more stable than the flow of one, but this has not proved to be the case. Silver in the last half century or so has been produced in such abundance that had bimetallism been continued the value of the dollar would have depreciated greatly. There is a further difficulty, as we have shown by our sketch of the monetary history of the United States. It is not difficult to select a mint ratio which at any one time is approximately the same as the market ratio. But the market ratio depends upon the relative quantities of gold and silver coming from the mines and upon the changing demand for these metals. When either gold or silver increases sharply, because of discoveries of bonanza mines, the market ratio is likely to change. The mint ratio, being fixed by law, cannot change continuously, and the result is that the metal which is being produced more freely begins before long to drive the other out of circulation. This in itself is perhaps no evil; but the history of silver-mining indicates that there would have come with this substitution of actual silver monometallism for theoretical bimetallism a sharp rise in prices. Bimetallism, by shifting the standard always to the cheaper metal, has a bias in the direction of price inflation.

This brings us to an account of recent monetary history. On March 6, 1933, owing to a nation-wide run on banks caused by a general fear for their solvency, specie payments were suspended, and on March 20 an embargo was placed on the exportation of gold. Thus the gold standard which had prevailed since 1879 was temporarily abandoned. Our monetary unit, however, again was tied to gold on January 31, 1934, when by Presidential proclamation issued under powers granted by the Gold Reserve Act of January 30, 1934, the gold dollar, formerly weighing 23.22 grains of fine gold, was reduced to 13.714 grains of fine gold, a reduction of approximately 41 per cent.

Under the Thomas Amendment of May 12, 1933, the President was empowered, at his discretion, to restore bimetallism. This power, however, has not been used. Instead, other legislation was passed which has again increased the importance of silver in our monetary system.

The Silver Purchase Act of June, 1934, declared it to be the policy of the United States ultimately to increase the silver monetary stocks to 25 per cent of the total value of the monetary gold and silver stocks combined. The Secretary of the Treasury was authorized and directed to buy silver in such quantities, under such terms, and at such times as he should deem reasonable and advantageous to the public interest. He could pay not more than 50 cents an ounce for silver held in the

United States on May 1, 1934, and not to exceed \$1.29 (the monetary value of silver) per ounce for other silver. The Secretary was directed to issue silver certificates in an amount not less than the cost of the silver purchased or more than its monetary value.

On the basis of the 1937 gold stocks (\$12,000,000,000) it would be necessary to acquire a total stock of \$4,000,000,000 of silver to reach the accumulation aimed at. At the time of the passage of the act the monetary stocks of silver amounted to about \$800,000,000. By March, 1937, the stock of silver dollars, silver bullion, and subsidiary silver had risen to about \$1,700,000,000.

Under the silver-purchase program, the rapid rise in the price of silver caused an appreciation of the Chinese monetary unit. This produced deflation and monetary disturbances in China and finally led that country to abandon the silver standard.

DEPOSIT CURRENCY

In modern society money constitutes only a small part of our medium of exchange. In fact no more than 6 per cent of our trade is carried on by means of money; the other 94 per cent of our buying and selling operations are effected through money substitutes. The most important of these are the check, the bank draft, and the money order. Most purchases made in modern society are paid for by check. We have now to inquire into the origin and nature of these money substitutes.

A check is an order drawn upon a bank to pay money to the bearer to whom it has been properly made out or indorsed. The person who writes the check is the *drawer*, or *maker*, of the check, the bank upon which it is drawn is the *drawee*, and the person to whom the money is to be paid is the *payee*. Now it is obvious that a person cannot draw checks on a bank unless he has a deposit at the bank which entitles him to order the bank to pay out money (currency).

This brings us to a consideration of the origin and nature of a deposit. The word implies that something has been deposited at the bank. You may go to your bank with a bag full of money and deposit it there, receiving a credit for the amount so deposited. This credit constitutes a *deposit account*. It consists of a promise on the part of the bank to pay out money on demand, up to the amount of the deposit, as ordered by you through the drawing of checks. Now it is clear that if all demand deposits originated in this fashion, they would be virtually warehouse

receipts for money, and not in reality a money substitute at all. In fact, only a very small part of our demand deposits originate through the actual deposit of money.

At the close of a day's business a mail-order merchant may have received, in payment for goods sold, a bundle of checks, bank drafts, and money orders. He probably has received little or no money. He takes these checks, drafts, and money orders to his bank and deposits them. In a sense he may be said to sell them to the bank, and the bank pays him in return by crediting him with their value on his deposit account. The bank promises to pay money on demand in return for the checks, drafts, and money orders so deposited.

CREDIT INSTRUMENTS

In large part, however, demand deposits originate in ways other than the two referred to above. A manufacturer or a wholesaler sells an order of goods to a retailer. The retailer is unable to pay cash for it. He gives the wholesaler or the manufacturer his promissory note for sixty or ninety days. The manufacturer or the wholesaler takes his promissory note to the bank, endorses it, and "sells" it to the bank. The bank pays him the face value of the note minus the discount for sixty or ninety days. Does the bank pay him actual money? Not at all. It credits him on his deposit account. Thus demand deposits originate in large part through loans.

It may be that the manufacturer agreed to carry the retailer on his books for sixty or ninety days. But the manufacturer is in need of circulating media to pay his obligations. He therefore borrows at the bank on his own promissory note for sixty or ninety days. The bank discounts the note; that is, credits him on his deposit account with the face value of the note minus the discount.

There is still another way in which the sale to the retailer may be financed. The wholesaler, upon having sold the goods, may draw a draft on the retailer for sixty or ninety days. This draft, or bill of exchange, is an order drawn on the retailer to pay the sum of money in question. He then takes it to his bank and has it discounted. At the end of the sixty or ninety days the banker presents it to the retailer for payment.

Deposits, then, do not originate merely through the deposit of money, checks, bank drafts, or money orders; they arise through the process of discounting. The bank "buys" promissory notes or bills of exchange and gives in exchange its own promise to pay money on demand. But

the bank's promise to pay money on demand (bank credit) is, in practice, as good as money itself, since checks properly made out are usually accepted as media of exchange.

THE VOLUME OF DEPOSIT CURRENCY

Demand deposits are frequently referred to as "deposit currency." The total volume of deposit currency in the country at any one time may be said to be measured by the total volume of demand deposits. The volume of *checks* in circulation represents the volume of deposit currency *in use* at any given time. Only a small part of the total volume of deposit currency is in use at any one time, just as only a small part of the money in circulation is in use at any one time.

As a rule, every bank carries a deposit account with each of several other banks. These banks are known as its correspondents. Banks may draw upon such deposits just as individuals draw upon their checking accounts. They do so by writing bank drafts. If you owed a debt to somebody in New York to whom your check would not be acceptable, you probably would go to your bank and buy a bank draft and mail it to your creditor. Bank drafts therefore are really equivalent to checks, since both are drawn against deposit accounts. Deposit currency thus circulates in the form of checks or bank drafts.

Bank deposits, it will be remembered, are promises on the part of banks to pay money on demand. At first thought it may appear that the total volume of such deposits would never be large because they would soon be paid out in money on demand. It may seem surprising therefore that the volume of demand deposits is, year in and year out, approximately five times the quantity of money in circulation. Deposits are surely intended for use. Why are they not drawn upon? The answer is that they *are* constantly being drawn upon, but that the checks are re-deposited as fast as they are drawn. Indeed, some checks are always being presented for actual money, but the money soon gets into the hands of someone who redeposits it. Promissory notes and bills of exchange are constantly maturing, but others are constantly being discounted. It is clear that deposit currency constitutes a medium of exchange which is constantly in circulation side by side with money itself. To be sure, there are seasonal and cyclical fluctuations in the volume of outstanding deposit currency and money. At times relatively more of the total circulating media consists of deposit currency; at other times relatively more consists of money. Thus during the fall of

the year, in the crop-moving season, there is a heavy demand for actual currency. At such times deposits are being converted into money.

Banks keep far too little money in their vaults actually to pay the total volume of deposits if all deposit claims were presented for payment at any one time. They must, however, keep a certain amount of cash in reserve, since a small proportion of the deposits (as noted above) is constantly being demanded in money. This accounts for the fact that such a large proportion of the total money supply of the country is not in general circulation, but is held by the banks.

We have now reached a point where it becomes necessary to describe briefly the structure of banking institutions in order to see clearly the interrelations of cash reserves, money in general circulation, deposits, and loans.

THE FEDERAL RESERVE SYSTEM

At the head of the banking system of the United States are the twelve Federal reserve banks. To these are attached the member banks, consisting of all the national banks and many of the state banks and trust companies. The Federal reserve banks are essentially bankers' banks. Their deposits consist almost wholly of obligations owing to member banks and to the United States government. The member banks serve the public; the Federal reserve banks serve the member banks. The Federal reserve banks are required to keep a cash reserve of 35 per cent of lawful money against their deposits; the member banks are required to keep, at the lower limit, a minimum reserve of 7, 10, or 13 per cent against their deposits, depending on which of three groups of cities a bank may be located in. Some fifty of the larger cities are in the 10 per cent class. The smaller places are in the 7 per cent class, and only banks located in New York and Chicago are in the 13 per cent class. These minimum reserve requirements may be raised 100 per cent by order of the Board of Governors of the Federal Reserve System. In the summer of 1936 they were so raised by 50 per cent, and in early 1937 to the full amount permitted by law. These actions raised the reserve requirements to 14, 20, and 26 per cent.

The member banks are required to keep all their legal reserves with the Federal reserve banks. Of course, they necessarily keep a considerable amount of cash on hand to meet their daily requirements. But the legally required reserve must be deposited with the Federal reserve banks in their districts. These constitute the bulk of the

deposits of the Federal reserve banks, against which, it will be remembered, only a 35 per cent cash reserve is required. Looking at the banking system as a whole, therefore, we see that only \$3.50¹ of actual cash need be available in the vaults of the banks to guarantee the redemption of every \$10 of deposits owned by the member banks, which in turn is the basis for the \$50 to \$100 (depending upon what reserves are required by the Board of Governors) of deposits owned by the public. As a rule, however, the Federal reserve banks hold about twice the cash reserves legally required.

Individual business firms, as we have seen, may present their promissory notes and bills of exchange for discount, and in return they ordinarily receive deposit credit. Member banks, in turn, may present this same commercial paper, if it meets the legal requirements, for rediscount at the Federal reserve bank in that district. When member banks rediscount their commercial paper in this manner with the Federal reserve bank, they ordinarily receive in return either a *deposit* at the Federal reserve bank or *Federal reserve notes*. If the member bank wishes to increase its reserve in order to expand its loans, it will probably take the deposit account; if, on the other hand, the member bank needs more money to pay out to its customers, it will be likely to take the Federal reserve notes. It makes very little difference to the Federal reserve bank which is preferred. Federal reserve notes are a promise of the Federal reserve bank issuing them, just as are also the deposits. The Federal reserve bank must maintain a gold reserve² of 40 per cent against these notes and, in addition, an amount of commercial paper or government obligations sufficient to equal, together with the gold reserves, not less than 100 per cent of the notes issued. It is therefore slightly more advantageous for the Federal reserve bank to credit the member bank with a deposit, since a reserve of only 35 per cent is required in the case of deposits, and the reserve may

¹ Roughly the minimum reserves maintained by the member banks formerly amounted to about \$10 (actually the figure for March, 1936, averaged \$8.40 for all banks) for every \$100 of demand deposits issued by them; and, since reserves of the member banks appear on the books of the Federal reserve bank as deposits against which a 35 per cent reserve of lawful money is required, it is clear that a minimum of \$3.50 must be held against each \$10 reserve deposit, which, in turn, formerly represented \$100 of demand deposits issued by member banks. When, however, the reserve ratio required by the Board of Governors within the limits fixed by law was doubled, \$10 of member-bank reserves could sustain only \$50 of deposits held by the public in member banks.

² This gold reserve is now in the form of gold certificates, since the gold itself, in accordance with the Gold Reserve Act of 1934, is held in the United States Treasury.

be in any form of lawful money. Since, however, the vast bulk of the lawful money held by the Federal reserve banks consists of gold certificates anyway, the difference between these requirements is not great. It is clear therefore that the Federal reserve notes are closely analogous to deposit accounts with the Federal reserve banks. But only member banks, in the main, have deposits with the Federal reserve banks, whereas the Federal reserve notes rapidly get into general circulation. The deposits in Federal reserve banks get into general circulation only through the "deposit currency" issued by the member banks, which deposit currency is based on the reserves carried by member banks with Federal reserve banks. Since these "reserves" are in reality deposit accounts issued by the Federal reserve banks to member banks, it is clear that the deposit currency in general circulation is based on the "deposit reserves" issued by the Federal reserve banks. It should be noted, however, that a large part of the reserve maintained by the member banks with the Federal reserve banks was originally established by the actual transfer of lawful money.

Suppose, now, that the depositors in member banks are calling for cash, as at crop-moving time or during the Christmas season. At such times depositors tend to convert a part of their deposits into cash. How can the member banks supply this cash? They in turn can demand that such surplus deposits (over and above their legal reserve requirements) as are carried with the Federal reserve banks be converted into Federal reserve notes. The cash reserves of the Federal reserve banks will scarcely be affected at all by this demand for cash. Their deposit obligations have been converted into note obligations. In this respect the banking system of the United States is not nearly so flexible as the Canadian system, which permits every bank to do within limits just what the Federal reserve banks alone can do; namely, to convert deposit obligations into note obligations. In the United States member banks cannot pay off deposits by the issue of their own notes. But the fact that the Federal reserve banks can do so gives a large measure of elasticity to the system. Member banks may secure additional cash by rediscounting additional commercial paper, or by borrowing from the Federal reserve, using any of their assets as security for the loan,¹ and thus obtain Federal reserve notes with which to pay their deposi-

¹Member banks who have not sufficient paper eligible for rediscount may obtain advances from the Federal reserve banks on the basis of their own notes secured by any of their assets which are acceptable to the Federal reserve bank. The interest on such advances is, however, $\frac{1}{2}$ per cent higher than the rediscount rate on eligible paper.

tors. Thus member bank credit (individual deposits) is converted into reserve bank credit (Federal reserve notes). Indirectly, then, this system provides an elasticity similar to that of Canada.

THE MONETARY SYSTEM

The circulating media thus consist chiefly of deposit currency and notes. Deposit currency composes the greater part of the *circulating media*, and Federal reserve notes the major portion of the *money* in circulation. Behind deposit currency lie the cash reserves of the banks. In the case of banks which are members of the Federal reserve system these reserves are deposited with the Federal reserve banks, and behind these deposits in turn is the 35 per cent cash reserve in lawful money held by Federal reserve banks. Behind the Federal reserve notes is the 40 per cent gold reserve held by the Federal reserve banks. Much of the lawful money therefore is held as reserve against other money—for example, the gold certificates that are held against Federal reserve notes. But more is held as a reserve against deposit currency. A dollar of lawful money held in reserve may, under the lower minimum reserve requirements, become \$30 in deposit currency. Therefore money proper cuts a much larger figure as a reserve against deposit currency than as a circulating medium. The reserve function of money is far more important than the circulating-medium function.

Gold is the base of our monetary system. It is our measure of value. Upon this base is built a superstructure of other forms of money and of deposit currency. The size of the superstructure depends upon two things: the quantity of gold¹ and the quantity of money and deposit currency which the banks are capable of erecting upon each gold dollar without transgressing the limits of safety, which limits it is sought to establish by law.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "In the latter part of 1862 the gold in the gold dollar became worth \$1.15." How could this be possible? Explain precisely what is meant. How does Gresham's law apply to this situation?

2. If the market ratio in a bimetallic country was 17 to 1 and the mint ratio was 16 to 1, which metal would be driven out of monetary circulation?

¹ Under the Silver Purchase Act of 1934, the gold monetary stock may be supplemented by silver up to a ratio of 1 dollar of silver for each 3 dollars of gold. Silver is lawful also for reserve but not as a reserve against Federal reserve notes.

3. "As a practical banker I cannot agree that bank deposits increase as a result of an increase in loans. In our bank we are always anxious to increase our deposits, since this enables us to increase our loans." Evaluate this statement.

4. A Minneapolis flour mill (Company X) sells flour to a merchant (Mr. Brown) in Winona. Suppose settlement is made with a bill of exchange. Write out a facsimile of such a draft. Suppose settlement is made with a bank draft. Write out the substance of such a draft.

5. "Banks are peculiar in this respect, that they are the only business institutions that boast of the volume of their debts." Indicate what is meant by this statement.

6. "Banks issue money and money substitutes. But the quantity that they can issue is limited." Explain.

7. "Banks may be said to be institutions which manufacture money. But this right ought to be reserved exclusively to the government." Do you agree?

8. "It is a great social waste to use such a valuable commodity as gold for monetary purposes. Government paper money would serve just as well." Discuss.

9. About 6400 banks are members of the Federal Reserve System. These banks control two thirds of the banking resources of the country, but they constitute less than one third of the banks. Do you think that all the state banks should be compelled to become members of the Federal Reserve System?

CHAPTER XIX · Price Fluctuations



In this chapter we shall consider first the movements of prices historically, and then the effect of these price fluctuations on various classes in economic society. We leave to a subsequent chapter the consideration of the causes of these price movements.

RELATIVE PRICES, OR INDEX NUMBERS

The changes in the price of a certain commodity over a period of time may be noted by comparing the actual price in a certain month or year with the actual prices in subsequent and preceding months or years. If, however, we are interested primarily not in the actual price but in the *price changes* from month to month and from year to year, it will be preferable to use relative prices instead of actual prices. Relative prices show the percentage changes in prices from a certain base month or year. Thus, if the price in a certain year or month is 10 per cent higher than in the base year or month, and the relative, or index number, for the base year or month is 100, then the relative, or index number, would be 110. The device of relatives, or index numbers, makes it easier for the mind to comprehend quickly the degree to which the price in a certain period is above or below the price in another period.

Relatives, or index numbers, have another useful purpose: they make it possible for us to grasp readily the changes in the general level of prices of hundreds or even thousands of commodities. If one had before him the actual prices of several hundred commodities for a ten-year period, it would be absolutely impossible, even after extended study of the figures, to get any very definite idea of the extent to which prices generally had risen or fallen until these actual prices had been reduced to relatives, or index numbers. We shall now consider briefly how index numbers are constructed.

Table 22 gives the actual wholesale prices of five commodities from 1922 to 1935. If we reduce these actual prices to relatives, or index numbers, we obtain a table like Table 23. In this table the year 1926 is the base year with which all other years are compared.

TABLE 22. Wholesale Prices of Five Commodities

	FLOUR (AVERAGE PRICE PER BARREL)	MILK (AVERAGE PRICE IN CHICAGO PER 100 POUNDS)	BEEF (AVERAGE PRICE PER POUND)	MEN'S SHOES (AVERAGE PRICE PER PAIR)	ANTHRACITE COAL (AVERAGE PRICE PER TON)
1922 . .	\$6.96	\$2.372	\$.150	\$6.51	\$10.41
1923 . .	6.19	3.051	.158	6.43	10.87
1924 . .	6.95	3.023	.171	6.25	11.35
1925 . .	8.54	2.930	.180	6.39	11.11
1926 . .	8.15	2.940	.164	6.40	11.48
1927 . .	7.16	2.975	.186	6.43	10.96
1928 . .	6.90	2.735	.228	6.75	10.93
1929 . .	6.52	2.955	.231	6.75	10.88
1930 . .	5.34	2.920	.207	6.75	10.84
1931 . .	4.578	2.775	.156	6.66	10.86
1932 . .	3.98	2.18	.131	5.71	10.42
1933 . .	5.42	1.94	.095	5.44	9.67
1934 . .	6.77	2.26	.114	5.50	9.42
1935 . .	7.38	2.60	.176	5.50	9.42

TABLE 23. Index Numbers, or Relatives, of Wholesale Prices of Five Commodities (1926 = 100)

	FLOUR	MILK	BEEF	SHOES	COAL
1922 . .	85.4	80.7	91.4	101.6	90.6
1923 . .	76.0	103.8	96.3	100.4	94.6
1924 . .	85.3	102.8	104.0	97.7	98.8
1925 . .	104.9	99.7	109.8	99.9	96.8
1926 . .	100.0	100.0	100.0	100.0	100.0
1927 . .	87.8	101.2	113.0	100.5	95.4
1928 . .	84.7	97.9	138.8	105.5	95.2
1929 . .	80.0	105.8	140.7	105.5	94.7
1930 . .	65.6	104.5	125.9	105.5	94.4
1931 . .	53.0	99.3	95.1	104.0	94.5
1932 . .	48.9	78.1	79.5	99.4	90.8
1933 . .	66.6	69.5	57.9	94.8	84.2
1934 . .	83.1	80.9	69.5	95.8	82.1
1935 . .	90.6	80.9	107.1	95.8	82.5

From Table 23 we can readily tell at a glance the changes in price from year to year for each commodity; but how much had the prices of all five commodities fallen on the *average* in 1932, for example, below 1926? To answer this question it becomes necessary to construct an average index number. Now there are a great many different types of averages. Some give good results, and others give poor results.

THE MAKING OF INDEX NUMBERS

We shall consider only four types of averages: (1) simple arithmetical average, (2) weighted arithmetical average, (3) median average, and (4) simple geometrical average.

The simple arithmetical average is obtained by adding together the five index numbers (for the year 1932, for example) and dividing by 5. This type of average is defective in that it gives equal weight to each commodity, whereas they are, as a matter of fact, of unequal importance.

The weighted arithmetical average corrects this defect. It is calculated by multiplying each index number for any one year by an appropriate "weight," adding the results, and dividing by the sum of the "weights." If flour is relatively twice as important as milk, it is weighted twice as heavily as milk. The Bureau of Labor Statistics weights each of the 784 commodities from which it constructs its index number of wholesale prices according to the total value of each commodity sold in the markets. Each commodity thus affects the final index number in proportion to its relative importance.

TABLE 24. Average Relative Prices of Five Commodities, 1922-1935

	SIMPLE ARITHMETICAL AVERAGE	WEIGHTED ARITHMETICAL AVERAGE	MEDIAN AVERAGE	SIMPLE GEOMETRICAL AVERAGE
1922	89.9	87.9	90.6	89.7
1923	95.4	90.6	96.3	93.7
1924	97.7	96.2	98.8	97.5
1925	102.2	103.4	99.9	101.7
1926	100.0	100.0	100.0	100.0
1927	99.6	98.5	100.5	99.3
1928	104.4	103.3	97.9	103.0
1929	105.3	103.5	105.5	103.6
1930	99.2	94.8	104.5	97.0
1931	89.2	82.0	95.1	86.8
1932	79.3	72.3	79.5	77.2
1933	74.6	69.8	69.5	73.5
1934	82.3	79.7	82.1	81.8
1935	91.4	91.6	90.6	90.9

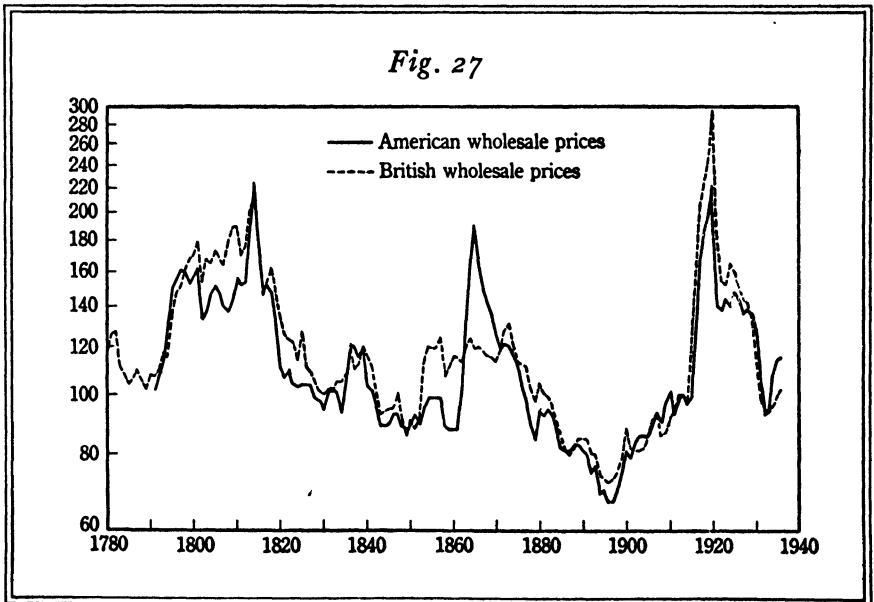
If no "weights" are obtainable, it has been found that the median and simple geometrical averages are to be preferred to the simple arithmetical average. The median is calculated by finding the index number which occupies a central position with reference to all the rest. It is

easy to calculate, since all one has to do is to arrange the various index numbers for any one year in order, from the lowest to the highest, and find the middle number. The simple geometrical average is the n th root of the product of the index numbers of n commodities. It is calculated by finding the logarithms of the various numbers for the year in question, summing these logarithms, dividing this sum by the number of logarithms summated, and then finding the antilogarithm of the number thus obtained.

These four types of averages are illustrated in Table 24. For the weighted arithmetical average the following "weights" were used: flour, 8; milk, 4; beef, 6; shoes, 1; coal, 5.

HISTORY OF PRICES

Fig. 27 shows the general movement of wholesale prices in England and the United States from 1790 to 1936. The curve for the United States is based on an average of about eighty commodities in the first part of



the period and over seven hundred commodities in the last part. The curve for 1790-1890 is constructed by the simple arithmetical method, and the weighted arithmetical average is used for the period from 1890 to 1936. The British curve is constructed by means of the simple

geometrical method from 1790 to 1850 and by the simple arithmetical method from 1850 to 1936. In spite of differences in commodities and method of construction the two curves show marked similarity, indicating that the forces affecting prices transcend national boundaries.

Let us center our attention on the American curve. Three large price upheavals are noticeable. The first was coincident with the War of 1812 and the Napoleonic wars in Europe, the second came with the Civil War, and the third with the World War. Wars have been major causes of price upheavals.

Certain broad trends are noticeable. The general trend was upward from 1790 to 1814. Then there was a long downward trend from 1814 to 1849; then an upward swing from 1849 to 1865,¹ followed by a long downward movement which culminated in 1897. This was followed by an upward movement until 1920. Then came a downward movement until 1933. Whether or not the upswing in recent years will prove the beginning of a long-run upward trend remains to be seen.

Superimposed on these long-run, secular trends are the short-run fluctuations of prices which correspond to the movements of the business cycle. From 1890 to 1915 there were five price waves of this sort. Thus prices are constantly rising and falling, but each succeeding wave may be higher or lower than the preceding, depending on whether the long-run trend is in an upward or downward direction.

EFFECT OF PRICE FLUCTUATIONS

Price fluctuations constitute one of the gravest evils of modern economic life. The mere fact of high or low prices is itself of no consequence. Sometimes we hear people talk as though high prices were inherently an evil and low prices inherently good. To be sure, everyone wishes to buy at low prices and to sell at high prices; but since buying and selling are two sides of the same shield, it is obvious that this is impossible. If prices run uniformly on a high level, incomes from property and labor also may be expected to run on a high level, and nobody is any worse off. Similarly, if prices are low, incomes from property and labor also will be low, and nobody will be any better off. High and low prices in themselves make no difference. It is the *change* from high prices to low prices or vice versa that hurts. Rising prices benefit certain economic groups and injure others; falling prices injure some and benefit others.

¹The high point was reached in England in 1873.

It therefore becomes necessary, in considering the economic effects of price fluctuations, to deal with economic groups or classes. Numerous groupings might be made, but we shall concern ourselves with two pairs of opposing groups of major importance from the standpoint of price fluctuations. The first pair are borrowers and lenders. They are especially affected by the long-run trends of prices. The second pair are the entrepreneurial class and the wage-earning class, who also are affected by the long-run trends, but more especially by the short-run cycles. Since the borrowing class is chiefly composed of entrepreneurs, we have here really three groups, entrepreneurs, capitalists, and wage-earners. Entrepreneurs are chiefly affected as borrowers by the long-run trends of prices, and as employers and enterprisers by the short-run price cycles.

How do the long-run trends of prices affect borrowers and lenders? Consider first the borrowing class. Let us suppose that a certain farmer bought a piece of land in 1864 at a price of \$20,000. He paid down \$10,000 and borrowed the remainder on a fifteen-year loan. In 1879 the wholesale prices of farm products had dropped to a point less than half as high as the prices that prevailed in 1864. Under these circumstances the farmer's money income was rapidly falling. With a declining money income it became increasingly difficult for him to pay the interest on his mortgage and to meet the payment of the principal when it fell due in 1879. Nor could he renew his loan of \$10,000 in 1879; for by that time the value of his land had fallen far below the 1864 level, and it would not have been safe for any investor to lend him three fourths or more of the value of his land. Not only was his income declining, but his original equity of \$10,000 in the land was vanishing. By 1879 he had lost a large part if not all of the money he had invested in the land. Falling prices were reducing both his income and his property.

Out of the decline in prices from 1865 to 1897 grew certain important social and political movements, such as the greenback, populist, and silver movements. Declining prices were robbing the borrowing class, and particularly the farmers, of income and property. No matter how hard the farmer worked and saved, the fall in prices defeated all his efforts. The dollar that he had borrowed had to be repaid in later years by a dollar that represented more products and therefore more labor and saving than the dollar he had borrowed. It was a "dishonest" dollar.

Consider, now, the effect of a rising price level on the borrowing

class. Suppose a farmer bought a farm in 1900 for which he paid, let us say, \$20,000. He advanced \$10,000 of the purchase price and borrowed the rest on a ten-year mortgage. From 1900 to 1910 the general level of farm-product prices increased 47 per cent. The farmer's money income was rapidly rising. With his enlarged money income it proved much easier to pay off the loan of \$10,000 than he had anticipated. He had borrowed dear dollars and was paying back cheap dollars. From 1900 to 1910 the value of farm land doubled. If our farmer was not disposed to save sufficient to pay off the loan, he would find no difficulty whatever in borrowing \$10,000 on a farm now worth perhaps \$40,000. Though he had not saved a cent of his yearly income, his equity in the land had increased from \$10,000 to \$30,000.

It is true that in the two periods which we have selected for illustrations there were other factors, besides the fluctuations in the general price level, which affected the farmer's income and property. The development of the factory system and large-scale production, together with the progressive cheapening of transportation, were bringing manufactured commodities to the farmer's door at more and more favorable prices. Likewise the disappearance of free land and the great growth of population from 1900 to 1910 accounts in part for the fact that the prices of farm products rose faster than general prices. This affected favorably not only the purchasing power of the farmer's income but the value of all his property. This aspect of the problem will be taken up in greater detail in a later chapter.

Turn now to the lending, or investing, class. Suppose a man left his widow in 1897 an estate of \$60,000 invested in bonds and mortgages. At 5 per cent interest this would yield her an income of \$3000. From 1897 to 1914 the cost of living advanced about 40 per cent. The \$3000 income would buy in 1914 only a little over two thirds of the goods and services that it would have bought in 1897. The rising price level had confiscated one third of the widow's real income. From 1914 to 1920 the cost of living more than doubled. The \$3000 income could buy in 1920 only about one third of the goods and services that it could have bought in 1897. The rising price level had in effect destroyed two thirds of the widow's real income.

A striking illustration of the effect of a rising price level is to be found in Germany's experience during the decade from 1913 to 1923. Suppose a man had invested 4,000,000 marks (approximately \$1,000,000) in high-grade bonds and mortgages. By August, 1923, prices had increased to such an extent that the purchasing power of

4,000,000 paper marks had been reduced to the equivalent of 6 prewar marks. Thus a million-dollar fortune invested in the safest possible manner had dwindled in ten short years to a paltry dollar and a half. The rising price level had confiscated practically the entire fortune. Subsequently prices rose still more until in November, 1923, it took a trillion paper marks to equal a gold mark.

Now it must not be thought that this loss was nobody's gain. The physical property represented by these bonds and mortgages was probably still as valuable in real purchasing power as ever. The owners of the property who had borrowed on bonds and mortgages were the gainers. The rising price level almost completely wiped out their entire debt. Thus fluctuations in the price level in effect take property from one class and hand it over to another class. Everyone recognizes the importance of legal protection of property rights, but few realize that fluctuations in the price level may reduce to an empty shell the property value ostensibly safeguarded by a bulwark of legal protection.

On the other hand, a falling price level favorably affects the investor. Every decrease in prices increases the purchasing power of his income and his property. His gain is the borrower's loss.

Consider now the short-run price cycles and their effect on the entrepreneurial and employing class on the one hand and the wage-earning class on the other. A period of rising prices is regularly a period of profit-making and prosperity for entrepreneurs, and a period of falling prices is a period of loss and depression. In the rising-price period all the entrepreneur's costs, with the single exception of raw materials, lag behind the prices of the products that he has to sell. But even on raw materials the entrepreneur often makes a profit in rising-price periods, since by the time they are worked up into finished products prices will have risen. This argument, however, would not apply to orders taken in advance of manufacture. Wages and all the fixed charges, such as interest, rents, and salaries, lag behind general wholesale prices. With prices rising and costs lagging behind, the margin of profit is widened. Moreover, a rising-price period comes with an expanding market. Thus profits increase for two reasons: because costs lag behind selling prices and because the volume of sales is increased. Conversely, in falling prices the entrepreneur suffers a loss; for sales fall off and costs remain high, whereas selling prices rapidly decline.

The wage-earning class is affected in two ways by the short-run movements of prices: as to employment and as to wage rates.

Employment rises and falls with the cycle of prices. Since the market is improved and the margin of profits is greater in rising-price periods, entrepreneurs are eager to increase production and employ more labor. Similarly, in falling-price periods, when the market is restricted and profits are dwindling and often losses are suffered, the entrepreneur reduces his force and sometimes closes down completely. Thus employment rises and falls with the movement of prices.

The effect of rising and falling prices varies with different classes of labor. For our purposes it is sufficient to distinguish between common labor and skilled labor. Employment increases for all classes of labor in rising-price periods and decreases when the price level is falling. But the fluctuations are much more marked in the case of common labor than of skilled labor. When an employer is contemplating a reduction of his working force, he naturally decides to lay off his common labor first. He is loath to discharge a skilled man who may be hard to replace when business picks up. The skilled men constitute the vital part of the working personnel, and it is important to keep this part of the organization intact as far as possible. Thus the skilled man is kept on the job, and he is kept busy partly by giving him unskilled work to do even though he is paid at skilled rates. Common labor is fairly homogeneous; one man is about as good as another, and any one is easily replaced. Hence the employer does not hesitate to discharge the common laborers. They are the first to go. Many skilled laborers suffer greatly from seasonal unemployment, but common labor takes the brunt of the unemployment which goes with price fluctuations.

Price fluctuations affect not only the employment of labor but also the purchasing power of its wages. Here we are interested in retail prices and the cost of living and not in wholesale prices. Retail prices and the cost of living do not fluctuate nearly as violently as do wholesale prices. Not only is the rise and fall less extreme, but there is a considerable lag. Hence even though wage rates regularly lag behind wholesale prices and fluctuate less violently than these, it does not necessarily follow that they fail to keep pace with the cost of living in rising-price periods or fall as rapidly in falling-price periods. However, wage rates frequently fluctuate less than the cost of living both on the upward and on the downward movement. Nevertheless, because of fluctuations in employment (both full-time and part-time) the total amount paid out in wages rises and falls more than does the cost of living. Thus labor as a class gains in real income in the upswing of the price cycle, and loses in the downswing.

Here, again, the effect varies with different labor groups. So far as skilled labor is concerned, it is usually true that wage rates fail to keep up with the cost of living in rising-price periods, but also fall less rapidly in declining-price periods. With unskilled labor the opposite is usually true: wage rates rise faster than the cost of living in the upward movement and fall faster in the downward movement.

Common laborers therefore are better off in the upward swing of the price cycle, both because of fuller employment and because of the higher purchasing power of wage rates. Likewise depression is an unmitigated evil for this group, both because of greater unemployment and because of reduced real wage rates. For the skilled laborers the case is not so clear. In the upward movement of the price cycle they suffer a distinct loss, since wage rates do not keep pace with the cost of living. On the other hand, they gain through fuller employment. In the downward movement they lose by diminished employment, but gain in so far as the cost of living falls more rapidly than wage rates.

So far as industrial unrest is concerned, the effect of price fluctuations is clear. A rise in price gives rise to unrest for three reasons: first, the menacing rise in the cost of living, which, as we have seen, particularly affects the more aggressive and better-organized laborers (the skilled); second, the growing volume of profits, which becomes generally known and leads to aggressive demands for higher wages; third, the increased demand for labor, which makes labor more independent and less wary of risking a struggle. On the other hand, falling prices also give rise to labor unrest. Wage cuts are resisted with widespread strikes, and unemployment gives rise to radical movements looking to a modification of the economic structure which makes such conditions possible.

Labor history is full of illustrations showing the influence of price fluctuations on industrial unrest. The first great labor movement in the United States grew out of the enormous upheaval in prices from 1833 to 1837. National unions sprang up with the price upheaval of the Civil War and disappeared almost wholly with the panic of 1873. The membership of the American Federation of Labor fluctuates with the movement of prices, gaining with rising prices and falling off in periods of declining prices.

Enough has been said to indicate that the problem of price fluctuations is one of the most important in the whole field of economics. We shall seek next to explain the causes of these price movements.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The real wealth of a country is not increased or diminished by fluctuations in prices. While these fluctuations are an injury to certain individuals, and a benefit to others, from a social point of view there is neither any gain nor any loss." Criticize.

2. "The price curve is the backbone of American history."—COMMONS. What is the significance of this statement?

3. "The depreciation of the monetary standard is a phenomenon of considerable social importance, and one whose effects must be regarded, on the whole, as beneficial. To begin with, its ordinary result is a rise in prices. Now a rise in prices is a useful stimulus to production: it keeps the spirit of enterprise alert; it encourages a rise of wages; it acts as a tonic; it is a symptom of sound economic health."—GIDE. Do you agree with this statement?

4. "If all prices and incomes rose equally, no harm would be done to anyone." Is this statement true?

5. "Before 1896 the 'bloated bondholder' was gaining. Money lenders like Russell Sage rolled up wealth. They could not have done so after 1896. Even had they saved every penny of interest and compounded it, they would have had less actual purchasing power now than when they started. The newly rich today are not bondholders, but stockholders."—FISHER. Explain.

6. "Falling as well as rising prices cause discontent. For example, before 1896 the western farmer hated the eastern capitalist whose mortgages he found increasing in weight owing, he thought, to some manipulation of the market of money or produce or both."—FISHER. Evaluate this statement.

CHAPTER XX · The Theory of Money and Prices

In the last chapter we discussed price fluctuations and their effect on various economic groups; in this chapter we shall seek to explain the underlying causes of these price fluctuations.

THE VALUE OF MONEY

To begin with, we should note clearly the difference between fluctuations in value and fluctuations in price. It is conceivable that the price of a commodity might change very greatly, yet its value might remain unchanged. The value of a commodity is the quantity of *other commodities* for which it will exchange; the price of a commodity is the quantity of *money* for which it will exchange. Now if the prices of all commodities exactly doubled, no change would take place in the value of these commodities, but prices would rise 100 per cent. Money alone would change in value. A unit of money now would buy only half as much as before. The value of a unit of money, like the value of any other commodity, depends on the quantity of goods it will buy. Thus the value of money varies inversely with the general level of prices; when prices rise, the value of money declines correspondingly; when prices fall, the value of money rises.

Standard money may itself be made from a commodity. Thus gold, which is our standard money, is itself a commodity. The value of standard money will always correspond to the value of the metal of which it is composed in a country which freely buys and sells that metal at a fixed ratio to the monetary unit. If the value of the 137.14 grains of gold metal (and remember that by value we mean the quantity of other goods that it will buy) is less than the value of a ten-dollar note, gold bullion will be presented at the Treasury for money in such units as are required by law, and vice versa. Thus in a country that is on the gold standard the *price* of gold never changes, but its *value* fluctuates inversely with prices. If the prices of all commodities (except gold) doubled, their value would be unaltered, but the value of the commodity gold would be reduced to half, while its *price* remained un-

changed. It is clear, then, that if *gold* is standard the money unit is tied to the commodity gold so far as its value is concerned.

In the case of limited coinage of gold it is conceivable that the standard money might consist of gold coin and yet gold might not be the standard of value at all. This could be true, however, only in case the coinage of gold were limited in quantity. In that case it is quite conceivable that a gold coin containing 100 grains of gold would have a value very much greater than 100 grains of gold sold in the market for manufacturers' uses. As a matter of fact, Austria at one time had as the standard money a silver coin of limited issue worth considerably more than the silver it contained. In this case silver was just a commodity like any other commodity, fluctuating in price along with other commodities as the monetary unit appreciated or depreciated in value. The fact that the money happened to be made of silver had apparently no effect on its value. What, then, did determine its value?

It has sometimes been contended that the value of money depends on its abundance or scarcity as compared with the quantity of goods for which it can be exchanged. Suppose the quantity of goods or commodities remains unchanged and the quantity of money doubles. Is it possible for the people of a nation to buy any more of the enjoyable things they desire simply because the quantity of money in circulation has doubled? Obviously not. It is clear that it is impossible to purchase any more goods and services than have been produced, regardless of how much money may be in the country; and if there are no more goods to be purchased than before, it is likely that more money will be exchanged for the same quantity of goods. This amounts to the same thing as saying that prices have gone up, or that the value of money has gone down.

Now it may have occurred to the reader to ask why a doubling of the quantity of money should cause any increase in the amount offered for goods. Might not the increase simply lie idle in banks, cash drawers, and private purses? Certainly, if the money were not offered in the market for goods, prices would not be affected. We therefore must distinguish sharply between the *actual quantity of money* in any given country at any one time and the *quantity of money offered in the market* during any given time. It is the latter that affects the value of money and the price level. When we speak of the "supply of money," we mean the quantity of money offered in the market in exchange for goods and services during any given time.

In the business world we say that we offer "money," even though we

actually present checks or drafts, which, as we have seen, are promises on the part of a bank to pay money on demand and are therefore as good as money. So when we speak of the quantity of money offered in the market in exchange for goods and services, we really mean the quantity of circulating media—deposit currency as well as actual currency.

THE EQUATION OF EXCHANGE

The quantity of circulating media offered in the market during a year in exchange for goods and services will depend on four things: (1) the quantity of money in circulation, (2) the number of times each unit of money in circulation is offered in the market during a year, (3) the quantity of deposit currency, and (4) the number of times the demand deposits have been drawn out and replaced. For these factors we shall use the following notation: M = the average quantity of money in circulation during the year (by "in circulation" we mean the money outside of banks and the reserves in the Treasury, and therefore available for the purchase of goods); V = the velocity of money, or the number of times each unit of money is offered for goods or services during a year; M' = the average quantity of demand deposits during the year; and V' = the velocity of demand deposits, that is, the number of times they are drawn out and renewed during the year.

Now, as we have seen in an earlier chapter, much of the money in circulation is based on reserves of gold or silver held in the vaults of the government and the banks. All the demand deposits are based on reserves of lawful money. Hence it is obvious that the quantities of M and M' are clearly related to the quantity of money held as reserves. If there is an addition to our stock of money either through an increased production or importation of gold or through an increased issue of government money lawful for reserves, this increase normally finds its way to the banks and thence to the Federal reserve banks, where it piles up as reserves.¹ Large reserves result in low interest rates (unless interest rates are consciously controlled), which in turn tend to stimulate borrowing at the banks. But an increase in bank loans, as we have seen, results in an expansion of demand deposits. Any large addition to bank reserves therefore will tend to reflect itself in an increase in M and M' . Every addition to the monetary stock has resulted in the long run (at any rate until recently) in a larger and larger

¹The gold must be deposited by the Federal reserve banks with the United States Treasury, whereupon gold certificates are received in exchange.

superstructure of circulating media built upon the foundation of these ever-expanding quantities of monetary reserves. In fact, deposit currency in the United States has increased much more rapidly than our monetary stock. This increase in the ratio of our deposit currency to our total monetary stock was due in part to the fact that an increasing proportion of our monetary stock was being held as bank reserves, but more especially to a more intensive utilization of our cash reserves. Thus the superstructure of bank credit was growing even more rapidly than the reserves upon which it was based.

All this amounts to nothing more than saying that loanable funds tend in the long run to be utilized. They may and often do lie idle temporarily in periods of depression, but the next swing of prosperity brings expansion. The monetary and credit facilities of a country usually find in the long run reasonably full employment. There are many outlets for the use of purchasing power, if not in commercial loans, then in investment loans. Under conditions of temporary or prolonged stagnation, additions to the monetary stock merely result in the piling up of large bank reserves unable to find an outlet in investment. Loanable funds and present purchasing power are then so abundant that there is no effective demand for them, and hence the funds lie idle in the banks. But historically it is evident that there has been at least a broad general tendency for loans and investments (and therefore bank deposits) to increase as bank reserves increase.

It is conceivable that if banking were conducted by a monopoly (or if the discount policy were closely controlled by a centralized body such as the Federal Reserve Board) loans and bank deposits might not increase in proportion to bank reserves. It might be argued that a banking monopoly could make bigger profits by restricting its loans and charging higher interest rates. But such a curtailment of loans and bank credit would tend to produce a decline in the price level; such a policy therefore would be of doubtful benefit to the banking monopolist.

Government regulation of banking policy might in the interest of the public welfare refuse to permit bank credit to continue to be the football of accidental variations in the monetary stock and the bank reserves. It is quite possible, therefore, that we shall not find in the future such a close relationship between the total monetary stock and the quantity of circulating media as we have in the past; but so far as these factors are permitted to work themselves out in an automatic self-regulating fashion, we may expect an increase in the monetary stock to reflect itself sooner or later in an increase in the circulating media.

V and V' depend largely upon (1) the density of population, (2) transportation facilities, and (3) the cyclical movement of business. If the population is becoming denser and the means for transportation more rapid and convenient, the velocity of the circulating media tends to increase. If business is good, money and deposits are being offered rapidly for goods; on the other hand, if business is depressed, money and deposits remain comparatively idle and are offered less rapidly in the market.

Prices rise, as we have seen, if an increased quantity of circulating media is offered in the market in exchange for a constant quantity of goods. But prices, or, conversely, the value of money, may be affected also by an increase in the volume of goods and services offered in the market. Suppose the volume of goods offered in exchange for money and its substitutes doubled while the quantity of circulating media offered in the market during the year remained constant, what would be the result? Clearly if double the quantity of goods and services were offered for the same quantity of circulating media, prices would fall to half the previous level unless the velocity of the circulating media had changed.

The quantity of goods and services offered in the market depends fundamentally upon two things: (1) the total volume of goods produced during the year and (2) the number of times these goods have changed hands in raw, semifinished, or finished form. The former varies with the productive capacity of society, the yield of the harvest, and the condition of trade; the latter depends chiefly upon the degree of specialization as between regions, industries, and occupations. If ten firms constitute so many links in a chain of production from the raw material up to the finished product, and each sells his product to the firm constituting the next step in production, it is clear that the volume of *trade* is increased thereby, even if there is no increase in the volume of *production*. The volume of trade we shall designate by the letter T .

As we have seen, prices tend to rise with every increase in the quantity of circulating media offered in the market, and to fall if the quantity offered decreases. On the other hand, prices tend to fall if there is an increase in the quantity of commodities (raw, semifinished, and finished) and services offered in the market, and to rise if there is a decrease in the offering of goods. Prices, therefore, vary directly with the quantity of circulating media offered, and inversely with the quantity of goods offered. We may express this relationship algebraically as follows:

$$\frac{MV + M'V'}{T} = P.$$

In this equation $MV + M'V'$ represents the quantity of circulating media offered, T represents the quantity of goods offered, and P represents the general level of prices. The relationship may be expressed more briefly by the following equation:

$$\frac{C}{T} = P,$$

in which C represents the volume of circulating media offered in the market and T the volume of trade. It cannot be emphasized too strongly that the volume of circulating media, C , is affected as much by V and V' as by M and M' .

The effect of C and T on the price level can be illustrated historically in the two periods 1873 to 1897 and 1897 to 1920. In the former period T was increasing far more rapidly than C , and hence the general level of prices was falling; in the latter period C was increasing more rapidly than T , and hence the general level of prices was rising.

THE QUANTITY THEORY

We are now in a position to consider a formal statement of a theory famous in economic literature—the quantity theory of money. John Stuart Mill, writing in 1848, put it as follows: "The value of money, other things being the same, varies inversely as its quantity, every increase of quantity lowering its value, and every diminution raising it, in a ratio exactly equivalent."

By "quantity of money" Mill means "the quantity of it which people are wanting to lay out; that is, all the money they have in their possession, except what they are hoarding, or at least keeping by them as a reserve for future contingencies." In other words, his terminology is equivalent to that which we have adopted above. "Supply of money" means the quantity of money offered in the market. "Money," says Mill, "acts upon prices in no other way than by being tendered in exchange for commodities. The demand which influences the price of commodities consists of the money offered for them."

He goes on to say that the "very same effect would be produced if we suppose the goods diminished, instead of the money increased; and the contrary effect if the goods were increased or the money diminished." And by goods being increased he means the goods offered in the market. Therefore he does not overlook the influence of trade on prices.

Nor does he overlook the influence of deposit currency. He says that

"general prices at any moment depend much more upon the state of credit than upon the quantity of money. For credit, though it is not productive power, is purchasing power; and a person who, having credit, avails himself of it in the purchase of goods, creates as much demand for the goods, and tends quite as much to raise demand for the goods, as if he made an equal amount of purchases with ready money. . . . All this purchasing power, however, is operative upon prices only according to the proportion of it which is used." In other words, Mill measures the amount of credit by the quantity of it actually offered in the market.

As thus stated, the quantity theory of money is absolutely incontrovertible. No one can deny the validity of the equation of exchange given above. It is a mathematical truism; for it must necessarily be true that the quantity of circulating media exchanged in the market must equal the quantity of goods exchanged multiplied by their prices. Thus $MV + M'V' = PT$. This equation merely points out a necessary relationship existing between six separate phenomena. A change in any one of them will necessarily result in a corresponding change in one or more of the other factors.¹ Let us suppose that P has increased. Then there must necessarily have been a corresponding change in some one or more of the other five factors. But which is cause and which is effect? At this point the quarrel over the "quantity theory" begins.

Let us narrow down the range of the controversy. We can reduce our six factors to four. M and M' may be lumped together under the one head of M , or circulating media. The ratio between the quantities of these kinds of circulating media varies with the customs and conveniences of the people and depends upon whether money or checks are preferred as means of payment. If we lump M and M' together, it is obvious that we must likewise combine V and V' and consider the velocity of the circulating media as a whole. $MV + M'V'$ thus becomes MV . Our four separate and distinct factors are then quantity of cir-

¹So far as goods are bought and sold on book credit, the two sides of the equation are not necessarily equal. If we may assume that the total value of book credit does not fluctuate materially from year to year, this factor is not significant. In the nineteen-twenties, however, it appears that there was a considerable expansion of book credit as a result of the increase in *installment buying*. Sales on the installment plan amounted in 1925 to about \$6,179,000,000, on which cash payments of \$1,638,000,000 were made, leaving \$4,541,000,000 to be paid on the installment plan. Over 60 per cent of the total value of these sales was for automobiles and trucks, nearly 30 per cent was for household goods, and only about 6 per cent was for clothing and jewelry. It should be noted that the dealer generally does not carry the customer himself, since, as a rule, he disposes of his installment paper to so-called finance companies or borrows more heavily at his bank. Thus installment purchasing is likely to give rise indirectly to an expansion of bank credit. Installment purchasing continued on a diminished scale during the depression, and is expanding again with business revival.

culating media (M), velocity of circulating media (V), volume of trade (T), and the general level of prices (P).

The quantity theorists argue that fluctuations in the quantity of circulating media (M) give rise to consequent movements in P . From this point of view it follows that price fluctuations are ultimately to be explained by changes in the relative quantity of circulating media.

THE INCOME THEORY

Albert Aftalion, a French economist, has recently advanced an income theory of prices.¹ Price fluctuations, he believes, depend upon the respective movements of the *money* income and of the *real* income, or, in other words, income measured in terms of commodities and services. If the *money* income of a given community increases while the *real* income remains the same, prices will rise; if the money income remains stationary while the real income is increased, prices will fall. This relationship may be stated by means of the equation $I = PO$, in which I is the money income, P the price level, and O the real income, or the output of goods and services.

The money income, I , may be broken up into Mv , in which M = the quantity of "money" (including deposit currency) and v = the circular velocity of money, or the number of times during the year in which money passes through a circuit of production from one income receiver to another. v must be distinguished sharply from V (see p. 332). The circular velocity (v) of money is much slower than the exchange velocity (V). The exchange velocity of money (also called "transactions velocity") is the number of times money is used in exchange. The circular velocity of money (also called "income velocity") is the number of times money becomes net income for someone during the year.

We may set off MV against transactions (the volume of trade, T), while Mv may be set off against output (the flow of real income, O). Thus we arrive at two types of price-level equations as follows:

1. The transactions equation

$$\frac{MV}{T} = P.$$

2. The income equation

$$\frac{Mv}{O} = P.$$

¹ *Revue d'économie politique*, May-June, 1925. Aftalion here carries forward a suggestion made by F. von Wieser in the *Schriften des Vereins für Sozialpolitik* in 1909.

The first shows the relations of (1) money, (2) exchange (or transactions) velocity, and (3) volume of trade, to prices; the second shows the relations of (1) money, (2) circular (or income) velocity, and (3) output, to prices.

The income theory has the advantage that it runs in terms of the modern theory of value with its demand-and-supply schedules. If the money income of a given economic society is increased, the demand schedules for goods rise, and prices are forced up.

The demand schedules for goods are in reality money-supply schedules; that is, offers of various amounts of money for different quantities of goods. As the money income increases, the amounts of money offered against given quantities of goods increase. This statement of the case is in no way contradictory to the quantity theory as explained above, but it has the advantage of stating the effect of increases or decreases in money income in terms of the supply-and-demand schedules for the numerous individual goods that are bought and sold in the market. Moreover, it views the problem less mechanically than the quantity theory does. According to the quantity theory we may expect that a certain increase in money income will result in a given increase in the general level of prices regardless of which sections of the community receive the increased income. But according to the income theory the increase in prices will be different, depending upon who the individuals are whose money income is increased. What classes of the people have received larger money incomes? rich or poor? savers or spenders? demanders of luxuries or demanders of necessities? The increase of incomes does not affect the price mechanically, or automatically, but through the desires of those whose incomes are raised.

Yet even here the income theory is not at variance with the quantity theory, properly understood. If the quantity theory is stated mechanically, and the assumption is made that with a given increase in the money supply everything else remains the same, it follows, of course, that the quantity theory, so stated, leads to wrong conclusions. In point of fact everything else does not remain the same. If the money income of the rich increases while the money income of the poorer class remains the same, shifts in demand for the different products will take place, and in consequence the relative production of different goods and services will be altered; if, on the other hand, the money income of the poorer class is increased while the money income of the rich remains the same, the shifts in demand will be of quite a different nature, and the subsequent shifts in production will be in favor of goods purchased by the poorer

class. Now, as we have seen, a change in production affects the price level just as surely as does a change in money income. Since the change in production is not the same when the income of the poorer class is increased as it is when the income of the rich is increased, it follows that a given increase in money income will not affect the general price level to the same extent in both cases. One of the main objections to the equation of exchange is that it is likely to mislead us into viewing the problem as though we were dealing with a static society. We are likely to assume that "other things remain the same," when in fact they do not. It is quite possible that the income formula gives a more dynamic approach to the problem of price fluctuations. Possibly, also, the income formula may help us to apply the equation of exchange with better insight into the relationships of cause and effect between the various factors which it takes into consideration.

LONG-RUN PRICE TRENDS

As has already been pointed out (see Fig. 27, p. 322), long-run price trends are revealed in the course of commodity prices during the last 150 years. Without going into too much detail let us center our attention particularly upon the movements of the last 60 years, though in part of the discussion we shall go back to the eighteen-forties. A downward trend developed from the early seventies to the late nineties, and an upward trend from the nineties to the World War.

How can we explain these price movements? There are two major explanations: one runs in terms of gold and the supply of money; the other in terms of investment activity.

The gold and monetary explanation is based on the quantity theory of money, which has already been discussed. According to this view, the fall in prices from 1873 to 1897 was due to a growing scarcity of gold and the consequent failure of the monetary supply to keep pace with the growing volume of trade. It is alleged that while the monetary gold stock of the world was increasing during this period, trade was increasing still faster. Statistical investigations reveal a general correlation between the long-run price movements and the movements of the "effective" monetary gold stock. The "effective" monetary gold stock is calculated by dividing the monetary gold stock in any given year by a general index of commodity production (which latter is assumed to be roughly representative of the volume of trade). From the middle nineties to 1914 the effective monetary gold stock was rising and so

also was the commodity price level. Gold, in this period, was increasing faster than trade.

Of course it is not gold alone which affects prices, but rather the supply of money, or the circulating media. As a matter of fact, the volume of circulating media erected on the gold base did not increase in precisely the same proportion as the gold stocks. During much of this period, silver was an important form of money in many countries, and paper money also was in use in varying quantities. Moreover, the volume of circulating media created on the gold base did not constitute, during the period in question, a fixed proportion of the gold stock. For these reasons the changes in the monetary gold stock do not give an accurate picture of the changes in the money supply as a whole. Fluctuations in the effective monetary gold stocks do not, therefore, tell the whole monetary story.

Nevertheless it cannot be denied that an important monetary fact in the last quarter of the nineteenth century was the relative shortage of gold compared with the growing output of goods and of trade. Similarly, from the middle nineties to the World War, gold was becoming superabundant in relation to output and trade. On these facts the quantity theory erects an explanation for the fall in prices in the first period and the rise in prices in the second period.

The investment theory offers a wholly different explanation for long-run price movements. Let us begin with the period from the middle forties to the early seventies. This quarter century was one of great investment activity, particularly in railroads. The building of the railroads, supplemented by related investment outlets in the heavy industries and in construction, created a great demand for investment funds. To meet this demand the normal flow of savings was quite inadequate. Business enterprise borrowed heavily at the banks. Additional circulating media were created, and so the gold reserves were utilized more intensively. Moreover everyone who had funds quickly put them to work. The velocity of the circulating media increased. Thus the favorable outlets for investment, opened up by technological developments, created an expanded demand for investment funds and called forth a larger supply of MV . From this point of view it was not the existence of an enlarged monetary supply which caused prices to rise. On the contrary, the enlarged money supply was called forth, at least in part, by the expansion of investment in durable or capital goods. The money supply adjusted itself, according to this view, to the requirements of the business community.

From the early seventies to the late nineties, however, investment, it is said, was relatively sluggish. There was going on, to be sure, a large amount of investment in all manner of new developments and in the extension of old ones. But the outlets were insufficient to absorb the large flow of savings which increasingly wealthy communities were offering in the investment market. Investment activity, being relatively stagnant, demanded relatively few funds at the banks. M was in consequence not growing rapidly, and V was sluggish. The slow growth of monetary funds was the result of a slowing down in the rate of real investment in fixed capital goods.

With the turn of the century came a new burst of investment opportunities. This was the age of electricity, of the gas engine, of automobiles, of road-building, of public utilities. Investment activity again exceeded by far the flow of individual savings. Bank credit was extended, velocity increased. MV rose more rapidly even than output. Prices moved upward. The monetary expansion, which made possible a rise in prices, was caused, it is said, by the increase in investment activity.

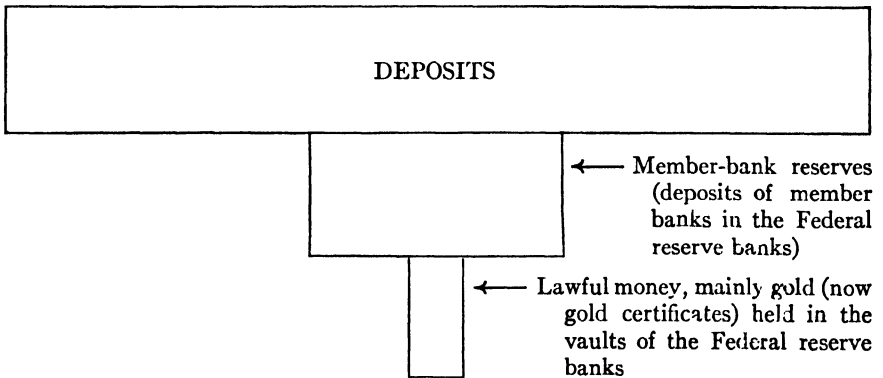
Both theories claim too much. Neither the money supply nor investment activity is wholly independent one of the other. Increased investment does tend to call forth an increased money supply (MV). But the money supply, in the period under discussion, was also in part a quite independent factor. The monetary gold stock was increasing at a rate considerably below that of the output of goods in the period 1873-1897, and considerably faster from 1897 to 1914. A relative shortage of money tends to check investment, while an abundant money supply tends to stimulate investment. Thus investment is in part dependent upon the money supply. On the other hand, investment activity is dependent not merely upon plentiful investment funds offered at attractive rates of interest; it is dependent also upon technological developments which open up attractive investment possibilities. Thus the two theories are not mutually exclusive, but in fact supplement each other.

PRICE MOVEMENTS SINCE 1914

The World War resulted in a great price inflation which ended in 1920. Then came a drastic fall in prices. From 1922 to 1929, prices remained nearly stable at about 40 per cent above the prewar level. In 1929 began a precipitous fall in prices which continued until April, 1933. From this point on, the movement was quite sharply upward until early

1935. During 1935 and until the autumn of 1936 prices remained fairly stable at an index of about 80 (that of 1926 being 100).

In order to discuss these price movements it becomes necessary to explain the extraordinary changes which have occurred in the place occupied by gold in the monetary system since 1913. To begin with, the introduction of the Federal Reserve System entirely altered the relation of the gold base to the monetary superstructure erected thereon. The establishment of the Federal reserve banks created a system of central banks. The reserves which banks were required to hold against their deposit liabilities were now no longer lawful money held in their own vaults, but simply deposits in the Federal reserve banks. These latter banks were indeed required to hold a reserve of lawful money. Thus gold—the chief form of lawful money—no longer served as a reserve against the deposit currency of the community. It became the reserve against these reserves, as explained on pages 314–317. The following diagram will indicate the new relationship.



The fairly close link that formerly existed between gold and the volume of deposits—between the gold base and the superstructure M —is thus severed. The Federal reserve banks can at will (within rather wide limits) increase and decrease the reserves of the member banks even though there is no change in the size of the gold base. This they can do through the sale or purchase in the open market of government obligations or of commercial paper. When the Federal reserve banks purchase such assets from the member banks the deposit accounts of the latter with the former are built up by the amount of the purchase. If the assets are purchased from the public, the Federal reserve banks pay by drafts drawn on themselves, which the public deposit with the member banks, who in turn send them in to the Federal reserve banks for

collection, thereby building up their deposits in these central banks. Conversely, the sale of government securities by the Federal reserve banks to member banks or to the public has the effect of reducing the balances of the member banks. Thus the reserves of the member banks may increase or decrease by reason of the open-market operations of the Federal reserve banks without any change in the gold base. Moreover, the Board of Governors of the Federal reserve bank may now vary the legal reserve ratios of the member banks. There is thus a very wide range of flexibility in the volume of M . The gold base does impose an ultimate limit, since the reserve banks themselves must hold a reserve of lawful money—mostly gold—of 35 per cent against the deposit liabilities. Since, however, they usually hold a reserve of twice this amount, a very large measure of freedom is accorded. The gold stock no longer determines in any such rigid manner as formerly the volume of the circulating media.

In addition, since 1914 the monetary gold stocks held in the United States have increased in a revolutionary manner owing in part to profound disturbances to the world's monetary systems occasioned by the war and postwar developments, and in part to the devaluation of the dollar in 1934. During the war, the belligerent countries abandoned the gold standard and adopted paper currencies. Gold was sent in large volume to the United States in payment for vast purchases of war materials. Thus the stock of monetary gold in the United States increased from 1.9 billion dollars in 1913 to 3.1 billion in 1919. Moreover, after 1920, owing in part to the shifts in trade and in international indebtedness caused by the war and postwar measures and in part to the disturbed monetary situation in foreign countries, gold continued to flow to the United States. Thus by 1929 the gold stock of the United States had increased to 4.3 billion dollars. It continued at about this level, with fluctuations in the interval, until 1934, when the dollar was devalued. When the gold content was changed in January, 1934, the monetary gold stock was automatically increased from 4 billion dollars to 6.8 billion. The gold stock continued to rise, partly from gold production, but mainly from gold importations, to 12 billion dollars in 1937.

Is there any relation between these gold movements and the movements of prices since 1914? Undoubtedly the violent upheaval in the world's distribution of gold, and the disturbed world situation which brought it about, had a profound effect on the war and postwar price movements. But the relation of gold to prices was never a direct or close one, as the following brief discussion will show.

The great inpouring of gold from 1914 to 1917 undoubtedly facilitated and in part caused the war inflation of prices in the United States. Without this net addition to the gold stock a monetary check would probably have restrained the price rise which developed before the United States entered the war. After the United States entered the war, a price inflation most likely would have been permitted to run its course with or without an adequate gold base. The presence of a large stock of gold enabled the United States to remain on the gold standard despite the great rise in prices. Prices, however, advanced to the full limit which the larger gold reserves made possible. The virtual exhaustion of the reserves of the Federal reserve banks forced a restriction of credit in 1920, whereupon prices collapsed.

During the decade of the twenties prices remained at a level far below the high point reached in 1920 but about 40 per cent above the 1913 level. Gold continued to flow in and was superabundant, so far as the United States was concerned, throughout the decade. But a large part of the world's monetary systems were left without an adequate gold base. Various measures were adopted to alleviate the situation. Gold was economized by the European countries by being taken out of general circulation and its monetary use limited to bank reserves and to international payments. Moreover, many countries adopted a gold-exchange standard whereby a large fraction of the reserves of the central banks could be held in the form of balances in gold-standard countries, thereby minimizing the use of gold stocks as reserves. These measures facilitated the general return to the gold standard by the various countries which, during the war, had been on a paper basis. The economy of gold effected in this manner made it possible to maintain, for some years, a gold-price level about 40 to 50 per cent above the prewar level. Had the world returned to the monetary systems prevailing before the war, it would not have been possible to sustain even for some years a price level appreciably higher than that of 1913.

It is difficult to state briefly the causes of the general world-wide collapse in prices which began in 1929. A full statement would require an analysis of the causes of the disastrous world depression. In general it may be said that the breakdown was caused by the numerous maladjustments incident to the war and to postwar developments. Among these the following may be listed: (1) the progressive accumulation of large agricultural surpluses, partly owing to technological developments and partly owing to the war stimulus to agricultural production in the newer countries, subsequently aggravated by the return of European

production (aided by high tariffs) following the war; (2) the war debts and reparations; (3) the abnormally large volume of postwar foreign loans and the shock caused by the sudden cessation of lending when the outlook for foreign investment changed; (4) the maldistribution of gold already referred to; (5) the return to gold by different countries at wrong foreign-exchange parities, the French franc being undervalued and the British pound being overvalued; (6) the failure of the international price system to adjust itself to the disequilibrium in the exchange rates, because of wage and other cost rigidities; (7) the unsound investments made by private individuals and by banks; (8) the speculative inflation of the American stock market with consequent overinvestment and misdirected investment.¹

Once the precipitous decline in prices was started, weak debtors were squeezed out. Forced sales of commodities and securities ensued, thus causing a further pressure on prices. An extraordinarily severe liquidation developed. In many countries the banking structure was too weak or too unsound to withstand the strain. In the United States many banks failed and others became insolvent, causing widespread distrust and the withdrawal of funds by depositors. This condition forced the banks to sell assets and call loans, which in turn resulted in further deflation of commodity and security prices. Finally in March, 1933, all the banks were closed, and only those which were found to be sound were reopened. The capital structures of those which could be saved were rebuilt, in large part with government funds, and a plan for the insurance of deposits was inaugurated.

The abandonment of the gold standard in April, 1933, and the devaluation of the dollar in January, 1934, directly affected the prices of raw materials which were sold in substantial volume in markets outside of the United States. Such commodities continued to sell in foreign markets at substantially the same prices in terms of foreign currencies. When, however, these currencies were converted into dollars, at the new depreciated exchange value of the dollar, the sums so converted amounted to more dollars. Hence the dollar prices of all such commodities rose. This applied to both export and import commodities.

Thus the depreciation of the dollar, in terms of foreign exchange, accounts very largely for the rise in American prices in 1933. In 1934 the main rise in prices was in agricultural commodities, and this rise was caused largely by the severe drought of that year. Subsequently

¹ Compare report by Alvin H. Hansen, *International Economic Relations*, pp. 122-123. University of Minnesota Press, 1934.

there was no important rise in prices until 1936, when another severe drought again drove agricultural prices up and the European rearmament affected the prices of certain raw materials. For the most part, the rise in the total national income in 1935 and 1936 was balanced by an increase in production—in real income—without any important rise in prices.

With respect to the future trend in prices, account must be taken of a wholly new fact in the monetary situation, caused by the depreciation or devaluation of most of the currencies of the world. This action, in effect, has enormously increased the monetary gold reserves. Not only has the increased price of gold stimulated gold production until it has reached a new high record in terms of weight, but each ounce of gold now counts as more monetary units. Thus the monetary value of the annual gold production has increased from \$400,000,000 in 1929 to \$1,200,000,000 in 1936.

All fear of gold shortage thus is definitely removed. Gold reserves are now amply sufficient to permit a substantial rise in prices in the next few decades. Is an upward trend in prices likely to develop?

On the basis of the quantity theory a secular rise in prices may be expected, provided the enlarged gold reserves are actually utilized. Formerly a fairly close relationship existed between the quantity of gold, the quantity of circulating media, and prices. As shown above, the direct connection between gold reserves and the superstructure of money no longer exists. Management by central banks (the Federal reserve banks in the United States) intervenes. The larger gold reserves merely give the central bank authorities greater limits of possible expansion. But these limits may not be exploited to the full. Nevertheless, since it is the tendency of the dominant elements in a competitive economy to favor monetary expansion, the abundance of the gold reserves is likely to provide relatively easy money conditions. In other words, the monetary conditions are likely to prove distinctly favorable to expansion—to a secular rise in prices.

But we must not overlook the important point made by the investment theory. Investment depends not merely upon favorable monetary conditions. It depends upon the prospective rate of profit on new investment. It is conceivable that the outlets for investment may prove relatively restricted, compared with such periods as 1845 to 1873 and 1897 to 1920. Will investment outlets develop equaling in importance (relative to the population and productive capacity of the current period) the railroads, the public utilities, the electrical-equipment in-

dustries, the automobile manufacture and highway construction, and the investment in new territories which former periods of great expansion provided? A general advance in housing standards, air-conditioning, and rural electrification suggest some of the possibilities. Unless there develop investment outlets adequate to absorb not only the flow of savings seeking investment but also a considerable volume of new bank credit, there will be no secular rise in prices.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It sometimes has been assumed that the quantity theory is obviously fallacious since we know from history that the quantity of money frequently has increased and yet there has been no increase in prices, and, conversely, the price level at times has changed without any corresponding change in the quantity of money. Similarly, it frequently has happened that the supply of a certain commodity has increased without any corresponding change in its value. Is it safe to conclude from this fact that the law of supply and demand has no validity? What possible factors have not been taken account of in both the cases referred to?

2. "But if we divest ourselves of all prejudices, and minimize the effects of local and temporary conditions by making long-range studies, in numerous countries, of price levels and money quantities, our search brings us to a conclusion from which there is no escape. We discover that there have been no large, sustained movements in prices without corresponding changes in the relation of the volume of money to the volume of goods."—FOSTER AND CATCHINGS. Illustrate from American economic history.

3. "If fresh means of payment are created—whether in the form of legal tender currency or in the form of banking means of payment—additional purchasing power is created. The result must be a rise in prices. We cannot say exactly how great this rise will be."—CASSEL. Why not?

4. "Between May of 1920 and December of 1924 the volume of monetary gold in the United States increased 73 per cent and the net demand deposits of reporting banks increased 14 per cent. Commodity prices at wholesale, however, dropped 37 per cent during this period. During the same period changes occurred in British money, credit and prices which the quantity theory cannot possibly explain."—ANDERSON. Do you agree? Why, or why not?

5. "Obviously there is a great shortage of money in the United States, since the property value of the country is about twenty times as great as the circulating media. Every dollar of wealth ought to be matched by a dollar of circulating media." Criticize.

6. "The greenbacks depreciated in value not because of the large quantity issued, but because of the loss of public confidence in the ability of the government to redeem the greenbacks." Evaluate this argument.

CHAPTER XXI · Business Cycles

INDUSTRIAL FLUCTUATIONS

The wholesale-price chart which appears on page 322 shows three types of price fluctuation: first, it reveals three periods of extraordinary price inflation; secondly, it shows unmistakably six long-run trends in prices—three upward trends culminating in 1814, 1865, and 1920, and three downward trends culminating in 1849, 1896, and 1933; thirdly, it shows numerous short-run fluctuations, which indicate the ups and downs of business prosperity and depression. It is with these short-run fluctuations that we are concerned in this chapter.

The term "business cycles" is of recent origin. Formerly economists and business men centered their attention on crises, panics, and depressions; they did not think of industry, commerce, finance, and business in terms of *cycles*. The increasing application of statistical research in the field of economics has revealed unmistakably that business does not normally run along on an even plane of prosperity only to be thrown suddenly from time to time into deep gulches of depression. On the contrary we find that business and industry are constantly rising and falling in more or less rhythmic waves, or cycles. Whether we look at the physical volume of production and sales or at the prices of economic goods, we observe this same cyclical rise and decline. These wavelike movements, or cycles, appear in every field of industry, commerce, and finance. They may be seen in the production of such basic materials as pig iron or coal, in the volume of construction work, in the nations' exports and imports, in the employment of wage-earners in factories, in the volume of manufacturing production, in the volume of freight carried by the railroads, in the prices of commodities sold at wholesale and retail, and in the changes in wage rates. On the stock market they appear in the prices of stocks and bonds and in the volume of sales of securities. In the field of finance they are evident in the movements of loans, deposits, reserves, discount rates, commercial-paper rates, call-loan rates, bank clearings, and liabilities of business failures. In fact the whole range of business activities runs in wavelike movements, or cycles.

Now it must not be thought that the mere fact that these different phenomena of economic life show up-and-down fluctuations constitutes in itself any proof that there is such a thing as a business cycle. If these fluctuations were merely sporadic oscillations back and forth without relation to each other, they would not constitute a business cycle. In reality, however, we find that the wave lengths of these various movements in any one complete cycle are substantially equal, and that there is close correlation between these fluctuations. This at once establishes a strong presumption that these various movements are all the result of a common set of forces which operate to produce what is known as the business cycle.

But while the fluctuations of these various component parts of a single cycle have a common wave length, they are not all synchronous. Some series respond very quickly to an upward and downward impulse; others lag behind in varying degrees. Commodity prices, for example, usually lag behind output, and wage rates behind employment.

VARIOUS PHASES OF THE BUSINESS CYCLE

The business cycle has several phases, which may be denominated as follows: (1) the prosperity phase, (2) the crisis phase, (3) the depression phase, and (4) the recovery phase. We shall now consider in some detail the chief characteristics of each of these phases.

In the prosperity phase of the cycle we usually observe the phenomena of rising prices. The selling prices of the finished products of manufacturers and the stocks of merchants rise more rapidly than costs. Wage rates do not start upward until some months after wholesale prices. Interest rates on current loans lag months and at times even years behind the movement of prices. All the overhead expenses of the business remain relatively fixed. Rents are generally fixed by long-term contracts. The interest charge on outstanding bonds remains fixed until the bonds mature, generally a period of ten, twenty, or thirty years. Salaries change very slowly in spite of a general upward movement of prices, and at times even run counter thereto. Freight rates and the rates on power and light, regulated as they are by public authorities, move slowly and tardily during the major portion of the upward movement. Although the prices of raw materials rise fully as rapidly and as high as the wholesale prices of the finished products, yet the producer who manufactures for stock enjoys a profitable spread between the prices paid for raw materials and the prices subsequently received for

the finished product. Manufacturers, retailers, and wholesalers find their inventories rising day by day in value. So the spread between costs and selling prices is widened. Profits rise, not only because of this favorable price margin, but also because of the increase in the volume of production and sales. New plants and additional equipment are installed. Retailers, wholesalers, and manufacturers, encouraged by rising prices and growing markets, increase their stock of goods. Expansion is the order of the day. Rising prices, lagging costs, a wide margin of profit, and large inventories are the characteristic features of the prosperity phase of the cycle.

Finally the time may come when factories are running at full capacity and no further increase in physical production is possible. The competitive bidding for goods and labor, however, continues, made possible by further borrowing at the banks and the consequent further increase in the volume of circulating media. In short, the purchasing power of the community measured in terms of money units continues to rise in spite of the relative limitation in the volume of goods available for purchase. Two results follow from this: prices continue to rise still further, and an increasing number of orders remain unfilled.

During this phase of the cycle the demand schedules of all sorts of buyers are rising. Buyers, especially in dealers' markets, are willing to pay more and more money for the *same* quantities. The demand schedules are moving upward, since the banks are furnishing, through the increase in loans and deposit credits, more and more purchasing power measured in terms of money units.

There is, however, a limit to the possible expansion of the community's purchasing power, measured in terms of money units. The reserves of the banks finally approach the danger point and ultimately the limits imposed by law. Fresh loans can no longer be made, and the expansion of bank credit comes to an end. A check is thus placed upon the competitive bidding for goods and services. The upward movement of prices is checked. Businessmen who hold large stocks of goods become nervous, and finally, fearing a fall in prices, throw their stocks on the market. A general selling scramble ensues, and prices drop.

It may be that the upward movement is checked before the final limitations imposed by the reserves of the banking system as a whole are reached. For example, in 1923, when the member banks had reached the limit of their own capacity to extend loans, they were reluctant to extend their credit further by rediscounting at the Federal reserve banks. Moreover, the officers of the Federal reserve banks and the

Federal Reserve Board were reluctant to encourage heavy rediscounting, since they feared another period of inflation such as was experienced in 1919 and 1920.

Other factors may impose a check on the expansion of bank credit. A bad harvest may make businessmen fearful of the future and cause them to hesitate to borrow further. Certain lines of business may be overdeveloped, with a consequent fall in profits. This development in certain industries affects the stability of interrelated industries and causes businessmen to lose confidence in the future. A war will have its effect, more or less extensive, depending upon its scope. All these factors affect the course of the cycle and produce marked differences between various cycles.

Gradually business swings into the crisis phase of the cycle. The rate of increase of expansion begins to decline. Prices first rise less rapidly and then begin to drop. This development gives the cost items, which got a slow start in the upward movement, a chance to catch up with selling prices. Wage rates continue to rise after the price movement has begun to slow down. Discount rates on commercial paper rise sharply. Extensions of plant and new equipment installed in the prosperity phase of the cycle and purchased at high prices now become a heavy load for the company to carry, for the fixed expenses entailed by these extensions must be met each year in spite of the new turn of affairs. Salaries, rents, and other overhead expenses have been pushed up late in the prosperity phase of the cycle. When the market begins to recede and prices decline, these heavy charges on business remain as a millstone around its neck. With costs remaining high and prices falling, the profit margin is converted into a loss margin. The great problem for businessmen becomes not so much to make profits, but rather to prevent too heavy losses. In fact, the first task confronting businessmen when they become aware that the crisis has really come is to maintain their solvency. Not having anticipated the crisis, they find themselves tied up with numerous contracts and commitments, with loans and obligations entered into under the expectations of sales and prices which cannot now be realized. They must pay their debts or risk being pushed into bankruptcy. Hence accumulated stocks are thrown on the market at whatever prices they will fetch, and this intensifies the fall in prices. The receipts are not nearly sufficient to meet obligations. Accordingly there ensues a pressure at the banks for new loans and extensions of time on past loans as they mature. Each businessman finds it impossible to collect from his debtors or to pay his own

obligations. Credit thus becomes "frozen"; debts are not paid when due. Gradually, through rigid economy and saving, the credits begin to thaw out. This is the process of "liquidation." The final phase of the liquidation process consists in forcing into bankruptcy those who are totally unable to liquidate, thus "thawing out" by legal pressure the hardest-frozen credits. This last phase of the liquidation process becomes most widespread after the crisis is passed and industry has passed into full depression.

When businessmen become thoroughly aware that their expectations are not going to be realized because of the decline in the market and the fall in prices, the first job that confronts them is to maintain their solvency. This is the outstanding characteristic of the crisis phase of the cycle. It is chiefly a financial problem. As business passes gradually into a period of more or less prolonged depression the problem of production and sales comes to the foreground. The manufacturer is confronted with the prospect of a falling market; each new quotation brings news of a further fall in prices. He is confronted with high costs on every hand—high wages, large fixed interest charges on past loans, high rents, high salaries, high taxes, high freight rates and public-utility charges. If he continues to operate, he may find his costs higher than his selling prices. At first thought it might seem that the proper solution is to cut the Gordian knot by closing his plant entirely. But unfortunately for him, although his gross income would be nil, his costs in large measure would still continue should he resort to this drastic action. He could, to be sure, eliminate his wage and salary bill, but he hesitates to discharge skilled, trusted, and responsible men; he must keep the essential part of the organization together. In no way can he get out from under the fixed interest charges, the rentals, and the taxes. Moreover, he has a considerable stock of raw materials on hand which are constantly falling in value. This stock at least must be cleaned out before it is feasible to close up shop. When the old inventories have been worked off, some plants close down completely for a few weeks or longer, others keep on running at reduced capacity.

Every economy is resorted to in the utilization of office supplies, plant equipment, power, light, and materials, and every wage-earner or clerical worker who is not highly essential is discharged. Wages and salaries are cut down, though (except for unskilled labor) not in proportion to the fall in selling prices.

Merchants are faced with a situation fairly similar to that confronting manufacturers. When a drop in prices comes, they are faced with

the problem of what to do with their stocks of goods bought at high prices. At first they endeavor to work off as much of their stocks as possible at the old prices. Finally certain merchants, desperately hard pressed for cash, cut prices considerably in order to liquidate a material part of their stocks. Others, seeing what is coming, cut prices in order to maintain the physical volume of sales. Eventually new stocks come on the market at lower prices, and all retailers are forced to reduce prices. In the face of this declining market, merchants, like manufacturers, endeavor in every possible way to push down costs.

As business proceeds into the depression phase of the cycle, bank loans and deposit credits are reduced, and reserves accumulate. The loaning power of banks thus becomes more and more favorable as depression develops. Yet during this phase of the cycle businessmen do not care to avail themselves of these credit facilities. The mere fact that the banks are in a position to extend loans is not sufficient in itself to cause an upward movement. Businessmen do not care to borrow at the banks unless they see before them prospects of making profits; nor do banks wish to extend loans except under similar conditions. The maladjustments of costs and selling prices have to be righted, and the excessive stocks of retailers, wholesalers, and manufacturers have to be worked off before the market conditions again become favorable.

Finally, however, recovery begins, partly by reason of increased investment in fixed capital, owing to the need for replacement caused by depreciation and obsolescence, and partly by reason of increased investment in working capital and inventory stocks. The continuous volume of purchases on the part of consumers eventually depletes existing stocks. At last certain retailers are forced to place new orders. For a time these orders are limited to immediate needs, for the retailers may still anticipate a further drop in prices. Gradually, however, as the orders accumulate from all quarters, the prices begin to stiffen. Retailers, becoming convinced that the bottom of prices has been reached, place still heavier orders. Wholesalers increase their orders from manufacturers. So the movement is passed on from the retailers to the wholesalers, from the wholesalers to the manufacturers of finished products, from these in turn to the manufacturers of semi-finished products in all the various stages, and from these to the producers of raw materials, lumber, iron ore, coal, etc. Transportation has been stimulated all along the line.

This increased demand for goods, reaching back from the retailers to all the preceding stages of production, finds expression in the enlarged

Fig. 28

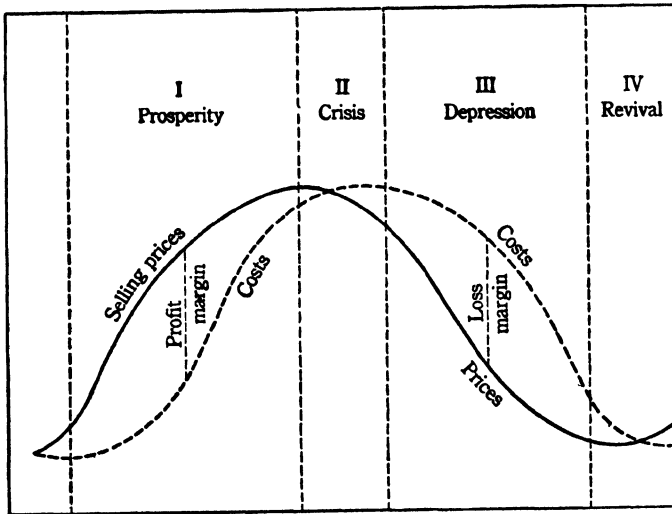


Fig. 28 gives a diagrammatic representation of the four phases of the cycle which are described in the text. The diagram shows the relation between costs and selling prices in the various phases of the cycle. The space between the two indicates the margin of profit per unit of output. During the first phase of the period of prosperity the margin of profit is widening, but during the last phase of the period of prosperity the margin is gradually narrowing, though still large. During the first phase of the prosperity period, therefore, the business outlook is steadily improving, and so the prices of corporation stocks are rising. When, however, the margin of profit begins to narrow, the business outlook is not so favorable and so the stock market begins to fall. Thus the stock market may reach the highest point before prices begin to fall and before the period of prosperity has come to a close. During the period of depression the profit margin is negative for many firms. During the last phase of the depression period, however, costs and selling prices are gradually approaching a more favorable adjustment, the future is beginning to look more hopeful, and so the stock market begins to move upward before the end of the depression period is reached.

purchasing power of all these dealers. Orders cannot be placed without the necessary means of payment. Only a limited increase in orders can be paid for out of the savings of past income. Loans are placed at the banks to finance these purchases. During the period of depression large reserves have accumulated at the banks, and as a result banks are eager to extend loans to sound businessmen at low rates of interest. The loans thus made are credited to the accounts of the retailers, wholesalers, and manufacturers in the form of demand deposits. Thus the purchasing power of all these businessmen measured in terms of money

units is increased, and with this enlarged purchasing power at their disposal fresh orders for goods are placed. Factories now begin to run at more nearly full capacity. More men are employed. Finally the competitive bidding for wage-earners results in higher wages. Thus the enlarged purchasing power at the disposal of the businessmen is finally passed on to the consuming public. Purchases in one field stimulate business in a whole list of related industries. With a larger volume of sales and a more favorable margin of profit the future looks brighter. Businessmen become optimistic. They stock up with raw materials, employ more labor, and enter into contracts for the construction of additional equipment and plant. All these activities in a highly inter-related industrial society have a far-reaching effect, and business rapidly passes on into the phase of full prosperity.

We now turn to the theories that have been advanced in explanation of these fluctuations in prices and production. We shall limit our attention to some of the more important of these theories.

RAINFALL AND BUSINESS CYCLES

William S. Jevons, an English economist of the latter half of the nineteenth century, and more recently Professor Henry L. Moore, of Columbia University, advanced the theory that the underlying ultimate cause of business cycles is to be found in cycles of rainfall. The causal relation runs, it is said, as follows: the cycles of rainfall produce cycles of crop yields; these cycles of crop yields produce cycles in the prices of agricultural products; these products in turn furnish three fourths of the raw materials of manufacture and thus profoundly affect the profit margin of manufacturers. Moreover, a heavy crop results in heavy traffic on the railroads. This not only results in large railroad revenues but forces the railroads to order new cars and equipment and to buy additional coal and other supplies, thus stimulating business activity in related industries.

Professor Moore contends that he finds evidence of an eight-year cycle in rainfall, crop yields, prices, and production. He holds that this eight-year cycle appears in the rainfall in Arizona, in North and South Dakota, in the Ohio valley, and in Illinois. The effect of these rainfall cycles is apparent on the growth of trees in Arizona; on the yield of crops in the United States, in the United Kingdom, and in France; on the wholesale prices in England; and finally on the production of coal and pig iron, basic industrial commodities. He finds that the maximum

points for all these cycles are substantially similar and occur in the following years: 1882, 1890, 1898, 1906, and 1914.

That crop fluctuations affect business conditions may be taken for granted, but that they are the dominating factor in the business cycle is more doubtful. In the first place, it is not clear that the length of the business cycle is really eight years. On the contrary, the length of the cycle has varied considerably during the last century. In the second place, the production of crops appears to oscillate back and forth without any well-defined cyclical movements such as would correspond closely to the business cycle. In short, the evidence as to the causal relation between rainfall and the business cycle is not sufficient to be convincing.

While economists generally do not have much confidence in the rainfall theory, it is certainly clear that oscillations of crop yield have important effects on the business cycle. A good harvest may initiate or intensify a period of prosperity. If timed properly it may hasten the coming of prosperity or postpone the period of depression.

SELF-GENERATING CYCLES

Professor Wesley C. Mitchell, director of research of the National Bureau of Economic Research, holds that the cycle is self-generating. Once started (and perfect stability in a complex society with a money and exchange economy is scarcely conceivable), it tends to perpetuate itself. Mitchell says that the cycle is not due to the influence of disturbing causes from the outside, but to processes which run regularly within the world of business itself. Prosperity becomes intense and breeds a crisis, a crisis runs on into depression, depression after a while begets a revival of activity, and the revival grows into prosperity. Thus the business cycles run an unceasing round, each cycle growing out of its predecessor and merging with its successor. For the explanation of business revival and prosperity it is not necessary to look to the occurrence of some propitious event, such as a good crop or the development of new industries (for example, the electrical or chemical industries) or a favorable piece of legislation or an increase in the export demand. "The quiet processes of business recuperation during dull times are quite competent to develop into revival without the adventitious help of any disturbing circumstances."

Why does one phase of the cycle generate subsequent phases? The answer is to be found in part in the accumulation, or storage, of stocks

of goods. In the period of prosperity and rising prices, stocks of goods are piled up for two reasons: (1) when prices are rising, retailers, wholesalers, and manufacturers all have a tendency to increase their stocks before prices rise still more; (2) in periods of prosperity the demand for goods is brisk, and so it becomes desirable to carry larger stocks to ensure an adequate supply. But of course retailers, wholesalers, and manufacturers cannot continue to increase their stocks indefinitely. As soon as they cease increasing their stocks (even though they continue to maintain them at the high level reached) the demand for goods at once declines in all markets except consumers' markets. Retailers buy less heavily from wholesalers, wholesalers less heavily from manufacturers, and manufacturers less heavily from the producers of semifinished and raw materials. Production is thus slowed down, wage-earners are thrown out of work, and this in turn reduces the demand in the consumers' markets. Prosperity lasts as long as stocks are piled up, but this development eventually produces a crisis and finally depression. At this point consumers' purchases begin to fall off and retailers become anxious to *reduce* their stocks. They therefore buy sparingly from wholesalers, supplying the consumers' demand from their accumulated stocks. Similarly, wholesalers will wish to reduce their stocks, and so they buy less from manufacturers than they sell to retailers. In like manner manufacturers will purchase less from the producers of semifinished goods and raw materials than they sell to wholesalers and retailers, since in large part they too are able to supply from their accumulated stocks such orders as come in. Thus the falling off in demand is increasingly greater the farther the market is removed from the ultimate consumer. Eventually the stocks are all used up; but in spite of depression and unemployment consumers continue to purchase at a fairly even pace, partly by utilizing personal credit at the stores and partly by drawing upon savings accumulated in the period of prosperity. Finally the stocks are largely used up. Retailers now are compelled to increase their purchases from wholesalers, and these in turn from manufacturers. In consequence of these increased purchases and the improved market conditions, dealers and manufacturers again begin to accumulate stocks. Retailers buy more from wholesalers and manufacturers than they sell to consumers, and manufacturers in turn buy more from the producers of semifinished goods and raw materials than they sell to wholesalers and retailers. Production is increased; more wage-earners are employed; and so the consumers' demand is reinforced and strengthened, and this demand in turn stimulates still

more the demand of retailers, wholesalers, and manufacturers. Thus the alternate accumulation and using up of stocks keeps the cycle moving forward from one phase to the next.

MONEY AND CREDIT

R. G. Hawtrey, a British economist, contends that trade fluctuations are due to disturbances in the available stock of money and bank credit. A contraction of currency occasions the slackening of demand for all classes of commodities, the fall of prices, the lack of employment, the shrinkage of profits, and the low rate of interest which are the concomitants of a trade depression. On the other hand, an increase of the currency leads to a stimulation of demand, a rise in prices, a high demand for labor, inflated profits, and a high rate of interest, all of which are the symptoms of active trade. A depression in trade is in essence a general slackening of the money demand for commodities; an expansion of trade is a general augmentation of the money demand for commodities. An expansion of trade occurs when the amount of credit money in existence is less than the bankers think prudent, having regard to their holdings of cash, so that they lower the rate of interest to encourage borrowing. Both dealers and producers borrow more money from the banks, credit money is increased, prices rise, wages are increased, cash is drawn out of the banks with which to pay wages, and bank reserves fall. At this point the bankers conclude that the amount of credit money in existence is more than they think prudent, having regard to their holdings of cash, and so they raise the rate of interest to reduce the excess. The rise in the rate of interest increases the cost of holding stocks of commodities, and dealers accordingly proceed to reduce their stocks by giving fewer orders to the producers. Production therefore is cut down, and wage-earners are thrown out of work.

What causes this expansion and contraction of currency? Hawtrey's answer is that in the banking institutions which prevail in the world as it is, there is an inherent tendency toward fluctuations. When credit money is curtailed, then business is depressed, unemployment ensues, and wages fall. The wages bill is the most important item in the community's demand for cash. If wages fell immediately as soon as banks began to curtail the credit expansion, cash would flow into the banks, the curtailment of credit would quickly be checked, and a stable condition might then be reached; but, instead, wages stay up for a considerable period, and when they eventually do fall and cash begins to flow

back to the banks, it is found that the credit contraction has been excessive since there now develops an overaccumulation of cash reserves. Eventually the banks must get rid of this excess, and the result is another period of inflation. Similarly, when the banks expand their loans, wages for a time fail to rise, and therefore cash is not drawn from the banks. Eventually when wages do rise and cash is withdrawn from the banks to meet the wages bill, the banks find that they have over-expanded, that their reserves are unduly low, and so curtailment again becomes necessary. Thus the cycle perpetuates itself.

PRICE FLUCTUATIONS AND THE BUSINESS CYCLE

Professor Irving Fisher has advanced the view that the fluctuations in the purchasing power of money largely predetermine the fluctuations in trade. It is generally recognized that a rising price level stimulates trade and that a falling price level depresses trade. In a rising-price period profits are increased (as we have already seen) because expenses, such as interest, rent, salaries, and wages, lag behind the rise in prices. Moreover, rising prices engender optimism, encourage speculation, and induce retailers, wholesalers, and manufacturers to stock up, not only because they anticipate a good market, but also because they anticipate a further rise in prices. As a result production is stimulated and industry expands. Falling prices, on the other hand, wipe out profits, force all dealers to reduce their stocks lest they be caught by a further fall in prices, and reduce the purchases of retailers and wholesalers; thus production is curtailed. Rising prices act as a stimulant to trade; falling prices depress trade. Fisher therefore contends that "changes in the price level afford an almost complete explanation of fluctuations in the volume of trade." This rise and fall in the price level, in turn, is controlled by extraneous forces, such as government inflation policies, banking policies, legislation, foreign trade, gold-mining, crops, and many other factors.

Fisher denies that the cycle is self-generating. He holds that Mitchell's view overlooks "friction." If we pull a twig and let it snap back, we set up a swaying movement back and forth. But the twig, once deflected and then left to itself, soon stops swaying. Friction brings it to rest. So in business we must assume that the effect of any initial disturbance would soon wear off after a very few oscillations of rapidly diminishing amplitude. The resultant business cycle would speedily cease altogether if dependent only on its own reactions. To keep it up, there must be applied some outside force.

OVERPRODUCTION AND UNDERCONSUMPTION

One of the oldest theories of the causes of depression, and one which still persists, is that of overproduction and underconsumption. The modern representative of this theory is John A. Hobson, a British economist. He argues that the ultimate cause of these cycles is to be found in maladjustments in the distribution of wealth. With the unequal distribution of wealth that now obtains there is an overaccumulation of savings relative to the volume of consumption. Wealthy people with very large incomes are unable or at least unwilling to spend a sufficient proportion of their incomes on consumers' goods to take the current production of such goods off the market. No matter how much production is directed toward the things that would appeal to the rich, whose wants, even the most ostentatious ones, are so amply provided for, they will inevitably invest too large a proportion of their incomes in capital equipment intended for further production. Since they themselves are unwilling or unable to spend the proper proportion of their incomes on consumers' goods, the investment made in new equipment must necessarily result in the production of goods intended for the masses. But the great bulk of the population, with their limited incomes, are unable to purchase all these goods. Thus production outruns consumption. The inevitable result is a congested market which leads to lower prices, diminished profits, and smaller incomes for the very rich. These reduced incomes force the wealthy classes to spend a greater proportion of their incomes to maintain their accustomed scale of living. Savings are thus reduced and less is invested in new capital goods. Consumption is thus given a chance to catch up with the productive capacity of the industrial equipment, surplus goods are worked off, and prices again rise. With the return of higher prices, profits are again increased, a disproportionate share of the national income again goes to the wealthy classes, and an excessive proportion of these surplus incomes is again saved and invested in new capital equipment which, of course, must be used finally in the production of consumers' goods. The masses, with disproportionately small incomes, are unable to purchase these goods, and so production again outruns consumption. Excessively large incomes are therefore responsible for the incessant attempt to employ capital in excess of the demand of the ultimate consumer.

According to Hobson, one of two things must be done. Either the rich will have to abandon their excessive saving and resort to greater spending, or else these surplus incomes must somehow be reduced and

greater equality of incomes attained. These latter ends could be obtained by the taxation of large incomes and by raising wages. A smaller proportion of the national income would therefore go to the rich, a smaller proportion of the national income would be saved, and thus a balance would be restored between consumption and production, between demand and supply.

It is to be noted that Mr. Hobson is not opposed to the accumulation of capital *per se*. He is opposed to saving a disproportionately large *percentage* of the national income; for that compels production to outrun consumption and results in these periodic stoppages, thus defeating its own ends. His point is that these excessive accumulations of capital, instead of increasing the total national income, actually diminish it, because of the unemployment that ensues. He wishes to have the accumulation of capital so nicely adjusted to the total national income that the consuming power of the community will always be able to take off the goods produced. If the distribution of income were more equal and the consuming power of society thereby enlarged, production, it is asserted, would become continuous. The total national income would then be greater; and even though a smaller *proportion* of the income were saved, the absolute volume of savings and capital accumulation might well be greater than now.

There are many known facts which *appear* to support this theory. It is well known that a large part of the savings of the community come out of the corporate surpluses retained in the business during the period of prosperity. Again, it is well recognized that no industry operates continuously at full capacity. This is true of the steel industry, the coal mines, and many other industries. They could not possibly dispose of their products at a profit if they produced continuously at full capacity. This would appear to give support to the notion that capital equipment is too extensive in proportion to national income. However, it is not difficult to account for this overexpansion of fixed capital. For one thing, when capital equipment is built it is necessary to make adequate provision for further growth and development. This factor alone accounts for much of the excess equipment. It would be uneconomical for any corporation to build a new factory just large enough to take care of its present business without making any provision for expansion. Moreover, it is exceedingly difficult to forecast the future accurately, and many errors of judgment are made. Part of the excess capacity can be justified by the prospects of growth, but a part must be charged up to unwarranted optimism. Moreover, a shift in demand

caused by the introduction of new products may leave an old industry, which faces a declining demand, oversupplied with fixed capital while the new one is undersupplied. Again, an industry may shift from one region to another because of a combination of circumstances. Thus the textile industry at present is developing rapidly in the South while mill towns in New England are declining, and flour mills are expanding in Buffalo while those in Minneapolis are operating at reduced capacity. Furthermore, new inventions make old equipment obsolete and account for a considerable amount of unused capacity. Finally, many industries are overbuilt in order to carry the peak load made necessary by the fact that we have a business cycle. If the business cycle could be eliminated, capital would be forced out of the overdeveloped industries and the consequent increased volume of capital seeking investment would cause interest rates to fall. Enterprises then could be started that now cannot be developed because the rate of interest is too high to make them profitable. It is not difficult to think of many schemes which are technically possible but which cannot earn the rate of interest that capital normally commands and which therefore remain undeveloped—power plants, reforestation, drainage schemes, irrigation, rural electrification, and so on interminably. It is to be doubted whether there is any end to the possible uses of capital if only it could be obtained at a sufficiently low rate.

The problem whether or not savings may accumulate in too large volume may mean one of two things. In the first place, it may mean that consumption is being curtailed too rigorously in the interest of provision for the future; that the present generation is being sacrificed in the interest of future generations. For example, a monopolistic corporation might be able to squeeze profits out of the public through controlled prices, or out of its wage-earners through abnormally low wages, the profits being used to finance construction requiring much capital and time for completion. Another example is the Russian communistic state which, it is said, has unduly curtailed current consumption in order to build up great capitalistic undertakings for the future. It has, moreover, been argued that the populations of the nineteenth century were induced, partly through preachments about the virtue of thrift, to devote too much productive energy to the accumulation of capital, thereby sacrificing current standards of consumption. The problem of the proper rate of capital accumulation is surely an important one, whether for a competitive economy or a state-controlled economy, but it has nothing inherently to do with the problem of the business cycle.

In the second place, the overaccumulation of savings may relate to quite a different problem. It may mean that the funds which are being saved are unable to find an outlet in investment in capital goods. The funds therefore are not used at all; they are "hoarded" in the form of idle cash or idle deposits, or are used to pay off loans at the bank, thus curtailing the flow of purchasing power actively functioning in the market. Such savings "run to waste." The funds are used neither for current consumption nor for investment, and so depression and unemployment ensue.

It is this last view which presents the essential problem involved in Hobson's theory. But Hobson's analysis fails to grapple effectively with it. The flow of savings, in fact, is relatively stable. It is the volume of investment in capital goods which fluctuates violently. The fluctuations in the level of investment cannot be explained by the *volume* of the flow of savings. Even though the volume of the flow of savings were sharply reduced in consequence, let us say, of a more equal distribution of income, the fluctuations in investment activity would still continue. At times the volume of investment would be so great that all available savings would be swallowed up and more—the balance being financed from bank credit. At times the investment activity would decline to such small proportions that no outlet could be found for the flow of savings. The real problem is: What causes these *fluctuations* in the rate of investment? Are these fluctuations independent of the volume of saving?

Should it turn out that the fluctuations in investment activity are essentially independent of the amount of saving, it would then follow that Hobson's proposed remedy—more equal income distribution—would not cure the business cycle, however desirable such a program might be on other grounds. This question is considered in the next section. Apart from the problem of cyclical fluctuations in investment is the *long-run* disequilibrium of investment and saving. This we have considered in the investment theory of price movements, discussed in Chapter XX.

INVESTMENT AND THE BUSINESS CYCLE

A long list of notable writers, including Spiethoff, Wicksell, Robertson, and Keynes, have advanced the view that the essential characteristic of the business cycle is the fluctuation in investment in capital goods, including (1) producers' goods (fixed equipment of all sorts),

(2) durable consumers' goods, such as houses and automobiles, and
(3) stocks of finished and unfinished goods. Put in a slightly different way, the essence of cyclical fluctuations is variations in the production of durable or capital goods. These writers differ considerably in their exposition, and we shall not attempt to give the precise views of any of them. What follows is a brief summary, on general lines, of this type of business-cycle theory.

The boom is characterized by a burst of investment. Investment activity is stimulated by a rise in the prospective rate of profit. These favorable prospects arise from far-reaching technological innovations, inventions, the discovery of new resources, the opening up of new territories, the development of new products. Not only are worn-out or obsolete capital goods being replaced; in addition there is a vast production of new capital goods.

The boom lasts so long as these new things are being built. But when they are completed, investment declines. You cannot satisfy the demand for a paved highway which will last, say ten years, by building one tenth of it now. You must build the whole highway. And when it is completed there is no need for another. So also with public utilities, electrification projects, the construction of apartment houses and office buildings, and the like. When these capital structures are completed there is need for maintenance and replacement, but, for some considerable period, not for much new construction. Even though there were no errors of overinvestment or misdirected investment, a decline in investment activity would still set in.

It is thus evident that because of the durability of capital goods it is highly probable that progress in the construction and accumulation of such goods will proceed by jerks. The period of prosperity lasts while the new capital goods are being constructed. The depression lasts until a large part of them have been worn out or rendered obsolete by technical progress, and until new developments create a demand for new types of capital goods. Fluctuations in investment activity arise from unequal rate of progress in a society using a large amount of capital goods.

The fluctuations in investment activity—boom and depression—are mainly the result of fluctuations in expectations about the future yield of capital goods. We have seen how technological developments open up favorable prospects for profit-making and give rise to an investment boom. These optimistic expectations are gradually undermined, owing to the increasing abundance of new capital goods (and consequently a tendency for their yield to fall) and also to the fact that

expansion increases the costs of production—wages, materials, and the rate of interest. Under these combined forces, disillusionment supplants optimism in business psychology. No one dares to invest in producers' goods or durable consumers' goods, or to accumulate stocks. Output and employment in the capital-goods industries therefore decline. This is the essence of depression. Business pessimism with respect to the prospective rate of profit is not easily dispelled. The depression thus wears on. How long it will take before expectations about the rate of profit are improved will depend upon the size of the previously accumulated stocks of goods, upon the length of life of the durable goods built in the preceding boom, upon how rapidly they are rendered obsolete by technical progress, and upon the speed with which new products are developed.

The cycle is thus mainly evident in the durable-goods industries: transportation structures and equipment, automobiles, building construction, machinery, and the iron and steel industries which serve these other fields. The cycle would persist even though consumer demand were completely stable, and even though the volume of saving were curtailed by more even income distribution. Since, however, a smaller flow of savings would limit the extent of the investment boom (unless indeed bank-credit inflation stepped in to fill the gap), the effect would be to reduce the *intensity* of the fluctuations. But this would be at the expense of progress to the extent that progress rests upon capital accumulation.

In this view the business cycle is a by-product of progress in a society which uses capital goods. The forces of progress—the discovery of new methods, new products, new resources—arouse bright hopes, engender optimism, and cause more fixed capital to be built. The results are a spurt of capital formation, an expansion of credit and of purchasing power, a rise in prices, a maladjustment of costs and selling prices, and finally depression as a result of these disturbed relations. Thus the business cycle is held to be a function of technological progress and its effect on the output of capital or durable goods.

THE BUSINESS CYCLE AN UNSOLVED PROBLEM

It is not possible, in the present state of economic knowledge, to give a thoroughly satisfactory explanation of the causes of the business cycle. To some extent it remains an unsolved riddle. Many economists are studying the problem, and it is one of the most fascinating though per-

plexing fields in economic theory. The apparent disagreement among economists, however, is not so great as at first appears. The literature on the subject is highly controversial, but the difference between the various writers is, to a large degree, one of emphasis.

There is one point to which we need to give attention with respect to these theories. It should not be forgotten that a description of the business cycle is a record of the actions of human beings. The business cycle is, of course, a product of human behavior. When it is said, for example, that the business cycle is a function of the money economy, what is really meant is that the money economy presents certain environmental stimuli to human behavior which cause actions such as are found in the various phases of the cycle. The same is true even of the rainfall-cycle theory. All these theories are really trying to explain a certain aspect of human behavior as we find it in the modern economic order.

DESIRABILITY OF ELIMINATION OF THE BUSINESS CYCLE

There are certain classes that make a living from speculation, and it is clear that the fluctuations of the business cycle increase the possibilities in this line of activity. These groups are naturally opposed to the elimination of the cycle. But many employers—perhaps most—would be glad to escape the uncertainties and worries that the cycle engenders. They wish to make their income from the production of goods, not from the uncertainties that the business cycle necessarily entails. Moreover, they would like to free themselves from the labor unrest that is caused by price fluctuations, by the lagging of wages behind the cost of living, and by the inevitable wage reductions that accompany price declines. For the wage-earner the cycle is a serious menace, since it is the fundamental cause of unemployment.

With all their evils, depressions serve to squeeze out inefficient firms, center attention upon ways and means of eliminating waste, and thereby prepare the way for a higher level of technical efficiency. Moreover, bold experimentation and new ventures are more willingly undertaken under the stimulus of boom conditions.

Since the business cycle is in large measure a function of progress in an economy dominated by private enterprise, it is doubtful whether it could be wholly eliminated in a free private-capitalist society. It is quite possible, however, that its violence may be mitigated. In any event, it is surely possible to alleviate its baneful consequences, particularly for the wage-earner, by social-security measures.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "There is a large and growing school of economists . . . who regard the trade cycle as essentially a monetary phenomenon, and in consequence expect to find the most efficient remedy in some modification of banking and currency policy."—BELLERBY. Do you agree? Why, or why not?

2. "The effective causes of trade fluctuations must be looked for in the sensitive and omnipresent machinery of money and credit."—HAWTREY. Argue for or against this statement.

3. Suppose a larger proportion of the national income were spent on consumers' goods and a smaller percentage saved, is there any reason to suppose that there would not still be fluctuations in prices? Suppose now such fluctuations occur. Would not all the phenomena of the business cycle—the lag of costs behind selling prices, the wide margin of profits, the piling up of corporate surpluses, the excessive accumulation of stocks, and finally the break in prices—at once appear? Is it likely then that the business cycle could be eliminated by more spending and less saving, by increasing the income that goes to the consuming masses, and reducing the incomes that go to the saving capitalists?

4. "In the depression there is a low rate of interest, and this has a restorative effect upon enterprise; in the high conjuncture there is a high rate of interest, and this acts as a brake."—CASSEL

5. It should not be concluded that a country can indefinitely prolong prosperity by the simple process of continuously inflating its currency. Such a procedure, it is true, may prolong the period of prosperity for several years. For example, during the years from 1920 to the middle of 1923, while other European countries were passing through a period of great depression and unemployment, Germany, with its constantly inflating currency and rising price level, showed many of the earmarks of prosperity, despite the political stress and turmoil of the period. Factories were running at relatively full capacity and wage-earners were fully employed. However, the rapid rise in prices placed a severe obstacle in the path of trade both internal and external, for no one knew what the monetary unit was worth from one hour to the next. Moreover, the sky-rocketing of prices so profoundly disarranged the distribution of wealth and income that industry found it virtually impossible to adjust its productive processes to the changing currents of purchases. The maladjustments created by the fearful upheaval in prices became so pronounced that the final collapse, which came during the latter half of 1923, was inevitable.

CHAPTER XXII · The Control of the Flow of Monetary Purchasing Power

COMPENSATORY CONTROLS VS. INDUSTRIAL REGIMENTATION

Broadly speaking, there are two types of control of the business cycle: (1) compensatory, or monetary, and (2) regimental, or industrial. Even if we hold that the causes of the cycle are nonmonetary in character, it is admitted that the effective supply of money is an important conditioning factor in both the upswing and the downswing. It is natural therefore that many should turn to the control of the money and credit mechanism as a means of tempering, if not removing, the excesses of booms and depressions. There are others, however, who have no faith in the monetary solutions. These believe that the problem must be attacked by each separate industry and by the co-ordination of the various industries. Stabilization is thus sought in terms of production control, not in terms of the control of money and credit.

The NRA and the AAA represented the industrial or regimental type of control. Open-market operations of the Federal reserve banks and a Federal budgetary policy are illustrations of monetary or compensatory controls.

The monetary type of control is of a general and indirect character. It does not submit each individual business firm or farmer to any specific regulation with respect to the conduct of his business. A high rate of interest imposed by central-bank action may indeed affect the business outlook. Such action influences the decisions of businessmen, but these decisions are made quite freely without any outside interference. Each entrepreneur is free to decide what he will produce, how much he will produce, and what prices he will charge for his product.

GOVERNMENTAL BUDGETARY POLICY AS DEPRESSION REMEDY

Government deficits are likely to arise in periods of depression, owing to (1) the decline of revenue resulting from business stagnation, and (2) emergency expenditures (*a*) for the relief of the unemployed and (*b*) for public works designed to provide employment. If these deficits

Control of Flow of Monetary Purchasing Power 369

are financed, either by the sale of bonds to individuals who otherwise would have held their funds idle, or by the sale of bonds to the banks, such deficit-financing results in the creation of new funds. These additional funds are then poured into the community through the payment of relief or through the construction of public works. The income paid out to these workers will be spent in large part in the purchase of goods which otherwise could not have been purchased, and this spending, in turn, creates employment and adds to the income of still other individuals. Such is the argument of those who advocate governmental expenditures as a remedy for depression.

A number of important qualifications must be made with respect to this sort of reasoning. Whether or not public expenditures will have the favorable effect indicated depends upon a number of conditions. To begin with, if the currency of a country is overvalued in relation to other currencies, so that costs and prices are too high in terms of the world market, such a country is likely to find the balance of international payments running against it, thus causing a loss of its gold reserves. If such a country insists on the maintenance of the gold value of its currency, as France did in the world depression, it must contract its income in order to deflate its costs and prices to the level of those of other countries. If it does not do so, gold will continue to flow out and it will be forced off the gold standard. Thus gold-standard countries with overvalued currencies are not free to engage in large public expenditures so long as they are unwilling to depreciate their currency.

In the second place, it is important to consider the possible effect of public expenditures, financed by borrowing, on the volume of private investment. If such expenditures have the effect of curtailing the volume of private investment by an equal amount, the net effect upon recovery is zero. Now it is true that during the depression, bank resources and other idle funds are so abundant that public expenditures are not likely to raise the cost of borrowing to private business, nor will the costs of labor or raw materials necessarily be raised thereby. Should such costs in fact be increased by reason of the public expenditures, private investment would thereby be discouraged. What is more likely is that private investment in large undertakings which reach far into the future will be deterred by a fear of future burdensome taxes or by a lack of confidence in the ability of the government to balance its budget within a reasonable number of years, thereby risking the precipitation of disastrous monetary disorders. If, however, the money and capital markets are favorable, and if the business community has con-

fidence in the government policy, the program of public expenditures is likely to be effective in enlarging the flow of purchasing power, increasing the total income of the community and thereby reducing unemployment.

It remains to consider the question whether such a program can in fact "prime the pump" so that with the cessation of the extraordinary public expenditures private industry can continue on its own momentum. It may be that the effect of expenditures such as those for public works is merely to maintain the level of business activity on a higher plane while the expenditures are going on and until the secondary favorable effects have spent their force.

There is little ground for believing that private enterprise will be able to carry on without the bolstering support of governmental expenditures until the time has arrived when the "natural forces" of revival have developed full strength. Not until depreciation, obsolescence, new products, and new developments again have opened up adequate outlets to absorb idle labor and plant in the output of capital goods can private industry go forward on "full steam."

On the basis of this analysis, it follows that public expenditure, under the favorable circumstances enumerated above, may prove valuable as a means of maintaining the level of employment and income on a higher plane during the depression, awaiting the time when private investment is able to go forward. Moreover, it may make the conditions for revival more favorable by facilitating the removal of surplus stocks of goods. As a means of "priming the pump," however, its efficacy is limited.

If public works are thrown into the market after revival has already got under way, the effect is likely to be an acceleration of the upswing movement. There is danger that this may lead to an excessive boom. Considering the functions which public works can perform best, such expenditures can be made most effectively during the depression, thereby decreasing its severity.

BANKING POLICY AND THE CYCLE

The Federal Reserve System has powers which in various ways may affect the flow of the national income. Certain measures based on these powers are designed to place a check on undue expansion; others, to check a deflation.

The expansion brakes may be enumerated as follows: (1) sale of government obligations in the open market, (2) increase in the re-

discount rate, (3) increase in the reserves requirements of the member banks, (4) "sterilization" of gold imports, (5) increase in the margin requirements on collateral loans on the stock market.

The sale of government obligations by the Federal reserve banks has the effect of tightening up the money market. This is true because such sales cause member banks to become indebted to the Federal reserve banks by the amount of such sales. The balances of the member banks with the Federal reserve banks are thereby reduced. Since these balances constitute the legal reserves of the member banks, the surplus reserves of member banks can be reduced by this method. If there are no surplus reserves, the action indicated will force the member banks to rediscount eligible commercial paper with the Federal reserve, or, lacking such paper, to borrow from the Federal reserve on their own notes secured by satisfactory assets. The Federal reserve banks, however, can restrain such action by raising the rate of discount. Thus the second check enumerated above may reinforce the first check.

The power to increase the reserve requirements has already been applied. On August 15, 1936, the reserve ratios were increased 50 per cent. This action had the effect of reducing the surplus reserves from about \$3,000,000,000 to about half that amount. The new reserve ratios against demand deposits were fixed at 10.5, 15.0, and 19.5 per cent, in place of the old 7, 10, and 13 per cent; for time deposits, 4.5 per cent in place of 3.0 per cent. By the beginning of 1937 the surplus reserves had risen again, by reason of gold imports, to over \$2,000,000,000. On January 30, 1937, the Board of Governors announced that a further increase in required reserves would be instituted, one half of the increase to be effective on March 1, 1937, and the other half on May 1, 1937. This action raised the ratios to 14, 20, and 26 per cent against demand deposits and 6 per cent against time deposits. It was estimated that this would eliminate about \$1,500,000,000 surplus reserves, thereby reducing surplus reserves to about \$500,000,000. Whereas in recent years the ratio of all deposits to reserves had averaged about 12 to 1, under the new requirements this became 6 to 1. The action taken did not in fact impose any restraint upon expansion, since it still left the banks with abundant reserves. It did, however, put the regulative machinery in a stronger position to put on the brakes by means of open-market sales of government obligations should it subsequently find it desirable to do so.

To prevent the continuing gold imports from creating new surplus reserves, the treasury, late in 1936, undertook to buy up and hold

in the Gold Stabilizing Fund all gold imports, the funds with which to buy this gold being obtained by the sale of Treasury obligations to the public or to banks.

CENTRALIZATION OF BANKING CONTROL

The Banking Act of 1935 moved farther in the direction of centralization of banking control. More definite responsibility is lodged in the Federal Reserve System to control the volume of credit in the respect that it is now equipped with more machinery for control than was formerly the case. Moreover, the development of the Federal Deposit Insurance Corporation has had the effect of extending the arm of Federal regulations over a larger part of the banking field. All members of the Federal Reserve System must belong to the Insurance Corporation. Until July, 1942, state nonmember banks may come under the insurance provisions, but after that date those with deposits of \$1,000,000 or more must become members of the Federal Reserve System in order to retain their insurance status. Thus the development of deposit insurance may help to extend the authority of the Federal Reserve System. There remain, however, side by side with the national banks with Federal charters and supervised by the Comptroller of the Currency, the state banks, operating under the jurisdiction of the forty-eight states. All national banks and some state banks are members of the Federal Reserve System. All national banks and most state banks are members of the Federal Deposit Insurance Corporation. Thus, while more banks have been brought under some form of Federal regulation, overlapping of control continues. In addition to the supervisory control of the state authorities in the case of the state banks, the Comptroller of the Currency with respect to the national banks, and the Board of Governors of the Federal Reserve System with respect to member banks, there is now added the supervisory control of the Federal Deposit Insurance Corporation.

The Insurance Corporation provides insurance for each depositor up to \$5000. Of the 15,800 banks in the United States about 14,300 are insured. Each insured bank must pay a semiannual assessment of $\frac{1}{24}$ of 1 per cent of its total deposits. This provides the funds from which deposit losses are to be reimbursed. The insurance covers about twenty billion dollars of deposits. Since the larger deposits are not covered this amounts to less than half the total deposits.

In consequence of the bank failures and bank reorganizations the

total number of banks has declined from over 30,000 in 1923 to about 15,000. The experience of England seems to point to the conclusion that a few large banks with numerous branches all over the country provide a more conservative, more responsible and safer banking structure than is possible with thousands of independent banking institutions. Branch banking is prohibited in many parts of the United States. Freedom to establish branches in all parts of the country would probably result in the evolution of a more integrated banking structure, better managed and better able to withstand financial strains. More important perhaps than any regulating machinery for the control of bank credit is the development of a sound banking structure. If the Federal Deposit Insurance Corporation succeeds in imposing higher standards of bank management, resulting in a higher quality of bank assets, it will perform a most useful function. There is some evidence that thus far it is contributing toward this desirable end.

QUALITATIVE VS. QUANTITATIVE CONTROL

Quality of bank assets is of necessity a matter of the highest importance. Nevertheless, it must be clear from what has been said above that quantitative control, within more or less flexible limits, has become not only a concern but a definite responsibility of central banks.

There are those who argue that qualitative control is all that is necessary, and, specifically, that it is all-important that commercial banks should keep their assets in so-called self-liquidating commercial paper, paper which matures in 30, 60, or 90 days, and which normally can be redeemed out of the proceeds of the sale of goods processed during the interval of the loan. There are others who argue, however, that banks are no longer able to find adequate outlets for the investment of their funds in self-liquidating commercial loans, and so are compelled to turn to bonds and mortgages. It is said that the modern large mercantile, manufacturing, or other industrial corporation finances itself mainly from the sale of securities and from its own undistributed profits, and so relies less heavily upon bank loans. Not only do the large manufacturing corporations not need loans for themselves; often they have sufficient funds to finance in large measure those who buy their products. Thus the characteristic function of commercial banks as the main agency financing working capital is said to be definitely on the decline.

Whether or not this will prove in the long run to be the case, at any rate it is clear that the postwar trend thus far has been in this direction.

If this trend continues, bank assets will be forced more and more into the investment fields.

This development does not necessarily mean a deterioration of bank assets from the standpoint of quality or even of liquidity. Obviously the bank must see to it that its assets are sound, whether they are in the form of commercial paper, mortgages, or bonds. In the case of real-estate mortgages it is important that long-term mortgages shall be amortized at a fairly rapid rate. Under the Banking Act of 1935 any real-estate mortgage loans which exceed five years must be amortized in accordance with specific requirements.

In order to secure liquidity, reliance in the past has been placed largely on short-term maturities. That this device is important so far as it goes cannot be doubted. Nevertheless it is increasingly recognized, partly as a result of the events of the world depression, that self-liquidating paper may turn out to be quite "frozen."

It is the first function of a commercial bank to extend loans—to provide its community with an adequate market for self-liquidating paper. But once these needs are taken care of, there is no good reason why banks may not invest in sound government obligations and in other soundly secured bonds and mortgages. It is important that a large proportion of these investments shall be in short-term maturities, and that even longer maturities shall be limited to a few years. In the case of mortgages adequate amortization is of utmost importance. In addition it is important that banks shall be in a position to sell sound assets, if necessary, to pay off their depositors. But if such sales are made in the open market the effect, if carried out on a large scale, is seriously deflationary, and may have disastrous consequences for the whole economy. It is therefore important that banks shall be in a position to sell assets to, or borrow on assets at, the Federal reserve bank. If a bank has eligible commercial paper it may rediscount it with the Federal reserve banks. If its supply of such paper is inadequate, it should be able to borrow on other sound collateral. This the Banking Act of 1935 permits at a rate of interest $\frac{1}{2}$ of 1 per cent higher than the rediscount rate. Such a provision helps to sustain the value of bank assets and hence the solvency of the banks when under pressure from depositors.

It is equally important to prevent undue pressure from depositors. So long as the Federal Deposit Insurance Corporation remains solvent, runs by depositors are not likely again to develop. The success of deposit insurance depends upon the solvency of the corporation, and this in turn depends upon the soundness of the insured banks. If the

corporation is not able to develop and maintain high standards of bank management in the participating banks it is inevitably doomed. Insurance is helpful to prevent the spreading of runs caused by the insolvency of a few banks. It cannot save the day if the banking system as a whole is unsound. Relentless insistence must be made on the quality of the assets of the insured banks.

Qualitative control is a first essential, without which quantitative control becomes quite meaningless. But it is equally true that unless the quantity of bank credit is kept within bounds, the quality of assets is endangered.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "If we could solve the 'money problem' we should have no need for NRA's and AAA's." What are the assumptions which lie behind this statement?

2. "Had the United States engaged in a large public-works program in 1932 the country would quickly have been driven off the gold standard." Argue for or against this statement.

3. "If a country with an overvalued currency undertakes a large internal expansionist program, increasing difficulties in its relation with the outside world will be encountered." Why?

4. "Public works are useful in a depression because they 'prime the pump' of private industry." Argue for or against this statement.

5. "Public works are of no avail as a depression remedy, for they only absorb capital funds which otherwise would be used by private industry." What points are overlooked in this statement? Under what conditions might the statement be true?

6. "Central-bank policy can stop a boom, but it cannot cure a depression." Is this statement wholly true? How would you rephrase it?

7. "It is said that easy money was a major factor in the revival of British industry from the depression, while in the United States governmental expenditures played a leading role." Why might a policy be useful in one country and not in the other?

8. "If we had branch banking in the United States we should not need deposit insurance." Argue for or against this statement.

9. "Under modern conditions, banks cannot find adequate outlets for their funds in commercial loans and hence are driven into investments." Point out tendencies which support or contradict this statement.

BOOK FOUR



THE DISTRIBUTION OF WEALTH AND INCOME

PART I. *Functional Distribution*

CHAPTER XXIII · Wages



THE MEANING OF DISTRIBUTION

With this chapter we begin a new division of the study of economics. In the preceding books we have examined the principles that control production, the pricing of particular goods, the value of money and the general level of prices, and fluctuations in industry. Now we turn to the study of the distribution of wealth and income.

The term "distribution" means the sharing of wealth and income among the people of the community. It does not mean (as it often does in common usage) the physical distributing of products to dealers or consumers from the places where they are manufactured or grown. In this and the succeeding chapters we shall be concerned with the forces that determine how the wealth and income of the country are shared among the people of the country.

A moment's reflection leads us to conclude that whether wealth is equally or unequally distributed depends in part on whether income has been equally or unequally shared among the people. For, although one may acquire great wealth by inheritance, the original accumulator must have enjoyed a large income. Of course wealth gives rise to income, and the process of amassing wealth is cumulative; having an income somewhat above the average, a thrifty person may accumulate wealth, and the possession of this wealth brings him a still larger income. But the fact still remains that inequalities in wealth are dependent on inequalities of income. It is logical, therefore, to treat of the distribution of income before that of wealth. The distribution of the total national income among individuals or among different groups of individuals is called the personal distribution of income. But, with certain exceptions, most persons receive their incomes from the control of one or more of the four factors of production. And whether any individual receives a large or a small income depends in part on how much wealth he owns. It depends also on the amount of income that is allotted to the factors of production he controls. If the returns to land are high, for example, landlords will receive large incomes; if wages

are high, laborers will have large incomes and probably a large share in the total national income. Therefore the personal distribution of incomes depends on (1) the distribution of wealth and (2) the absolute and relative amounts of income received by each of the four factors of production. It is necessary to understand the principles that govern the shares of the four factors before we can discuss with understanding the personal distribution of incomes. That phase of the theory of distribution which deals with these principles is called functional distribution.

THE IMPORTANCE OF DISTRIBUTION

The personal distribution of wealth and income is of the greatest importance in modern society. Many writers, following David Ricardo, have asserted that distribution is the principal problem of economics. Although this opinion may not be held universally, there is general admission of the vital relationship between the existing inequalities of wealth and income and many of the most difficult and important controversies in economic and political affairs. A few of these relationships may be mentioned to show concretely the bearing of inequalities in distribution on some of these problems.

The control of industry is largely in the hands of those who own the land and capital used in production. It is they who decide what shall be produced. They stand at the center of the industrial organization and make the decisions to buy or sell, to expand or contract production, and to make improvements or to retain old methods. But property-owners, of course, are not free to act unhampered by other groups who also have a role to play in production or consumption. The owners of the property are restricted by the laws and administrative decisions of the government, by the activities of trade-unions, by consumers' actions, and by many other social forces. Yet when all due allowance has been made for the effects of these opposing or limiting forces, it is still true that the owners of property are dominant in economic affairs. Because property-ownership is unequally distributed among the population, the major control of economic activity is vested in a relatively small number of persons. In a democracy this inequality of control gives rise to many political struggles. Having the power to control economic affairs, the property-owners have generally sought to strengthen their position by dominating legislation touching the ownership of property. The unpropertied classes also seek to dominate legislation for their benefit. Hence it is not surprising that in recent decades economic legislation

either for or against property and the income from it should have occasioned the most important political contests arising in times of peace.

The problems of distribution have a bearing on production. On the one hand, it is asserted that inequality in the distribution of income, which leaves a large part of the population with insufficient means for a comfortable standard of living, restricts production. It is alleged that if the poorer families had larger incomes they would be more efficient producers, and that in the next generation many of the children of these families would be able to advance still further in the scale of productive efficiency. On the other hand, those who oppose the intervention of the government for the enforcement of greater equality assert that if the rich had smaller incomes and the poor larger, less would be saved, and that the accumulations of capital would be smaller and the progress of civilization slower.

The principles of distribution are involved in many other problems and questions constantly before Congress, the state legislatures, and local authorities. The protective tariff has been held by its defenders to be the bulwark that protects American labor from the competition of the pauper labor of other countries. It is said either to cause the relatively higher wage level or to prevent its reduction. Other people assert that the recent restrictions on immigration are the first legislative acts that have actually protected American laborers from the competition of foreign laborers.

In the realm of private business also a knowledge of the principles of distribution is significant. The managers of public utilities and of other enterprises using large amounts of durable capital are constantly confronted with the question whether it is better to use labor or machinery to perform certain kinds of work. If wages are to rise in the future, some substitution of machinery for labor must take place; if they are not to rise, a heavy investment in such capital goods may prove unprofitable. A corporation that has a large bonded debt often must decide whether interest rates are likely to rise or fall in the future. If they are to rise, then it will be good policy, other things being equal, to issue long-time bonds; if they are to fall, it will be better to issue short-term obligations and refund when loans can be secured at lower rates.

THE EARNINGS OF LABOR

The functional share in distribution received by human beings in return for their services is wages. The workman who is employed by another person receives a definite contractual sum which no one would

fail to recognize by that name. The same is true of the earnings of all salaried employees. Though it is not customary to speak of the "wages" of self-employed artisans, businessmen, and professional workers, yet there is no fundamental difference between the contractual wage received by the common laborer in a factory and the fee of a physician, an architect, or a lawyer. All are payments for human service.

In many cases, however, there is great practical difficulty in distinguishing the earnings of labor from the receipts from other sources. A small storekeeper, a contractor, or a farmer always owns and employs some capital, and the last is frequently his own landlord. In all three cases the net income that the individual receives from his enterprise is a mixture of wages, capital return, and, in some instances, rent and profits. In larger businesses the separation between the earnings of the entrepreneur as a laborer and his receipts from property is not so difficult. This is owing partly to better accounting and partly to the fact that more than one person is likely to have an equity in the capital of the business, and, as a result, the managing entrepreneur will receive a stipulated salary in addition to whatever returns he may get as part owner of the business.

WAGES THE PRICE OF LABOR

Labor receives a payment for services just as the seller of any commodity receives a price for the things he offers to buyers. In some instances the service is purchased directly by the consumer, and the wage in that instance is determined according to the principles of value explained in Chapters VII-IX. When the supply of domestic servants is very large, their wages fall if the demand for them remains the same; if the demand increases, their wages rise. The same statement applies to the professions. The principle of demand and supply controls earnings.

There are, however, considerable differences between wages and other prices. If a merchant cannot sell his goods at the price he has set on them, he is likely to mark them down; but if a physician cannot get what he thinks is a proper fee, he may remain idle for a long time rather than accept work at a lower price. Moreover, although it is a daily occurrence for buyers to offer a lower price for a good than the seller asks—either because they expect to bargain a little or because they do not think the good is worth what the seller asks—there is less tendency for the buyers of direct labor services to attempt to bargain with the seller. If one thinks a physician's charges are too high, he

usually seeks out another whose rates are lower. When labor is employed by large firms, the employer sets a wage and the workers either accept it or look elsewhere for work ; or, if they do decide to bargain, in a majority of cases they will do so collectively through a trade-union.

Another difference between commodity prices and labor prices is the relative effect of custom. The prices of many sorts of goods respond slowly to changes in demand and supply, because of the fact that when a certain price has been paid for some time, neither buyers nor sellers readily think of offering or asking anything different. But in the case of wages this tendency is much more marked. No doubt the tendency is due in part to an ethical notion on the part of both employers and workers that the latter are entitled to the wage they have received in the past, because it is presumably a just and proper wage.

But in spite of all differences, wages are still the price paid for labor, and, like other prices, they must be explained in terms of demand and supply. The wages of those who render direct services are determined, as has been said, by the number of workers offering their work and by the demand prices of the buyers ; and the demand prices of the buyers are determined according to the principles of marginal utility. But in the case of workers who are demanded along with various kinds of capital, land, and entrepreneurship to make a common product, the determination of demand is more complicated. It is a relatively easy matter for one to decide whether he wishes to employ a mechanic to repair his motor car at a dollar and a half an hour ; he has only to consult his own satisfaction and his own means to find out whether the results will justify that expenditure. But it is far from simple for the manager of a plant to determine whether workers in a given class are worth thirty dollars a week. Their product is not easily isolated, as is the product of the mechanic. The remainder of this chapter will be concerned principally with the explanation of the demand for labor that does not work independently and has no separate product, but is one of several agents employed by entrepreneurs to work together to produce a common product.

THE MARGINAL PRODUCTIVITY OF LABOR

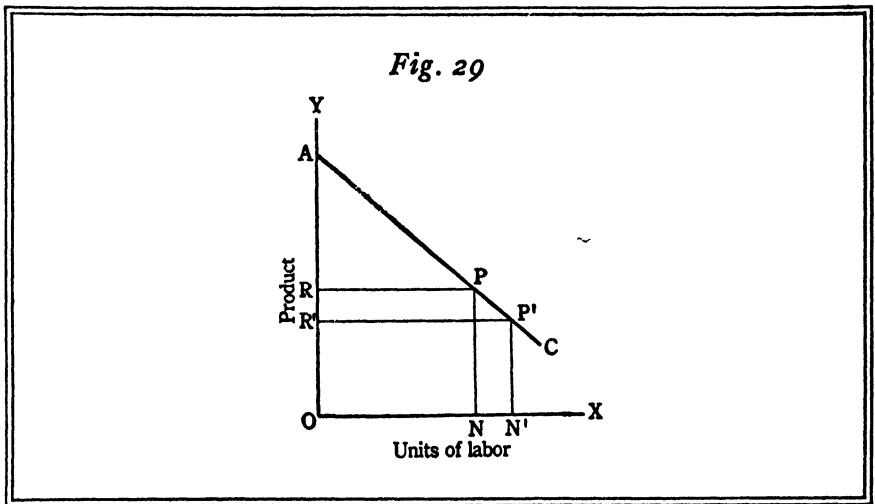
The typical employer's demand for labor depends upon the marginal productivity of labor. The laborer works with capital and land under the direction of the employer. Now the employer produces goods to sell, and he tries to produce them as cheaply as possible. To do this he

is constantly on the watch for methods of reducing his costs. One method of reducing costs is to substitute cheaper agents of production for those that are more expensive. Under conditions of perfect competition each entrepreneur, and under dynamic conditions the marginal, or typical, entrepreneur will select that combination of the agents of production which will bring his product to market at the lowest average unit cost. If wages are relatively low while capital costs are high, he will substitute labor for capital and increase his employment of workers; if, on the contrary, wages are high, he will try to put machines in the place of human hands and employ fewer workers.

We may suppose that the employer begins by adding labor to a given amount of capital and land. As he adds unit after unit of labor, the total returns from the factors will increase. At first, if the amount of capital and land is very great, the percentage increase in the total returns will be greater than the percentage increase in the number of labor units employed; that is, an increase of 10 per cent in the number of laborers will be accompanied by an increase of more than 10 per cent in the product. But long before the combination approaches the point of lowest unit cost the added product resulting from another laborer will have begun to decline. With wages at a given figure, the employer will add labor until the last unit just pays for itself. By the phrase "just pays for itself" we mean that the added product from the last unit of labor will sell in the market for just enough to pay the wage required to obtain that laborer's services. Here we may revert to the table on page 68 and the accompanying diagram. Labor as the variable factor will be increased to the point where the marginal cost of the product equals the selling price.

If we pass from the demand of a single employer to the demand of all the employers in society, the same explanation will hold. All employers will add labor up to the point where the last additional laborer adds only enough to compensate them for his wages; further than this they will not go. If all the workers who want jobs are not taken on when this point has been reached, there must be unemployment. And unless the unemployed are supported by their unions or by public funds, those who have no work will compete with those who have, and wages will tend to fall. The price of labor will continue to go down until all the workers can be employed. But, as we have seen, this point will be reached when the marginal workers throughout industry add just enough to the common product of labor, land, capital, and entrepreneurship to warrant their employment at the going wage.

To complete the explanation, assume that when the last available worker has been employed the wage paid him is less than the added product which accompanies his accession to the combination; that is, he and all other laborers are paid less than their marginal product. Under this condition the employers individually will fall short of the highest-profit combination; hence any one of them can increase the gains from his fixed factors of production if he raises wages a little and by this means adds to his labor force. But this is a resource open



to all employers, and all will use it. Competition therefore will force wages up to the point where the last unit of labor receives a wage equal to labor's marginal product.

In Fig. 29 the productivity of labor is represented by the curve AC. On the axis OX are measured the units of labor that are combined with the total amount of land, capital, and entrepreneurship in the nation. If ON units of labor are applied to the land, capital, and entrepreneurship in industry, the product of the marginal unit of labor will be NP. And since the units are interchangeable, every other unit of labor has the same marginal product. Each produces NP and each receives as its share, or wage, the money equivalent of NP. If the number of labor units is increased to ON', the marginal product of labor will decline to N'P', and wages will become the value equivalent of that product; if, on the other hand, the number of labor units falls to something less than ON, the marginal product will rise, and wages will be increased by a corresponding amount.

This statement of the general theory of wages necessarily makes certain assumptions. It assumes, in the first place, that the land, capital, and entrepreneurship are fixed in quantity, and that as variable amounts of labor are added to them the law of diminishing returns operates from the point where the smallest quantity of labor is applied to the fixed factors. The productivity curve, AC , is a curve of diminishing returns. In the second place, it is assumed that the labor is all employed. In the third place, the variable, which is measured along OX , is units of work and not human beings. If one man can do twice as much physical work in a day as another man, he is obviously putting forth twice as many *labor units* as the other. It is labor units, of course, that give rise to product, and it is the marginal labor unit that creates the marginal product of labor. This distinction between laborers and the work, or labor units, they put forth is of importance for more than logical precision. Laborers, as we know, vary considerably in their ability to do the same sort of work and also in their ability to do different sorts of work. Herein lies one of the chief reasons for individual as well as regional differences in wages.

In the foregoing explanation of the determination of wages we have assumed also free competition. We have supposed that workers compete with workers, that employers compete with employers. Trade-unions have been assumed to be absent, and understandings among employers have been left out of the picture. But even if we remove the assumption of free competition, the marginal analysis will not be overturned. Results, however, will be different from what they are under competitive distribution. If trade-unions limit the supply of workers in an employment, they may be able to force wages above the marginal productivity of the normal supply of labor within the industry. But they cannot hold them at this level unless the marginal productivity of labor also has been driven up through the limitation of the supply of labor. If they exclude laborers who might desire to enter the industry and who would find employment there if no union existed, then, by their restrictive organization, they increase the supply of workers in employments outside the union. Because the principle of marginal productivity prevails in these outside employments, and because the laborers excluded by the unions will not find capital in these employments increased as much as the labor force is increased, their marginal productivity will be less than it was in the employments from which they came. All consumers will be affected somewhat by the raising of the wages in the unionized industries. The entrepreneurs in these industries cannot

buy raw materials at lower prices or borrow capital at a lower rate than the other entrepreneurs in the market. The increased wages added to these normal costs of production will tend to raise the price of the product of the unionized industry. Of course, if only a small segment of the entire industry of a country is controlled by unions that practice exclusion and that force wages above the normal marginal productivity of labor, the burden on the consumer will be negligible.

In stating the theory of marginal productivity we have assumed that the entire labor force was employed when (Fig. 29) *ON* units of labor were combined with the land, capital, and entrepreneurship. If that is not the case, then the normal productivity of labor will be less than *NP*. There will be a tendency, if competition prevails, for wages to fall until all laborers have found work. However, if the proportion of the total population physically able to work was limited by a law forbidding young people to be gainfully employed until eighteen years of age, the normal productivity of labor would be the margin reached when all those legally employable had been absorbed into industry.

We have assumed also a perfect market. Laborers are supposed to know where they can sell their services to the best advantage; they are supposed to be free to move from place to place, from trade to trade, and from job to job. If they do not have this freedom, as is often the case, then we shall find that laborers are too numerous in some places, trades, and jobs, and too few in others. As a result, the marginal productivity of workers of equal efficiency will be unlike, and wages will be unequal. But in every case wages will approximate the productivity of the people employed.

It is sometimes said that an entrepreneur cannot discover what a laborer is worth to him, and that he knows only that he must pay the going wage; if his costs exceed the selling price he will fail in business, and if they are less he will make a profit. The complete answer to this objection must be deferred in part to the next section; here we need only to note that the entrepreneur *must* discover whether an additional laborer is worth his wage. To assert that he cannot discover when he has too much or too little labor leads to the absurd conclusion that any number of laborers might be employed with any amount of capital and land without affecting the amount of the product. It may readily be conceded that the adjustment of labor at the margin is inexact; only in pure theory are adjustments of any sort precise. Yet the inexactness of the entrepreneur's measurements is no reason for denying the operation of the general principle. The device by which he discovers the

productivity of labor is the common one of adding or subtracting a little labor and then noting the result. If the addition of this small amount of labor is accompanied by an increase in the product which is as great as the wage bill, then the laborer is worth hiring; if the increase is greater than the wage paid, then still more labor should be employed.

THE PRINCIPLE OF SUBSTITUTION

The entrepreneur, of course, does not carry on the experiment precisely as it has been described. In the first place, he may not regard labor in general as a factor of production. He may think in terms of different grades of labor and regard each grade as a separate agent of production. Instead of subtracting or adding labor to a fixed quantity of capital, he may add one grade of labor to a given quantity of capital and other labor. Moreover, he will probably contrast the profitableness of adding more labor with the profitableness of buying a machine, which amounts to combining more capital with the same amount of labor. If he finds that it is more profitable to add a machine instead of a number of laborers, then he will not hire additional labor; but it would not have been possible to add the machine and still improve the combination unless he had had too much labor for his capital equipment before he purchased the machine. In other words, at the rate of wages prevailing, the marginal product of labor was below the wage. By substituting capital for labor the entrepreneur raises the marginal product of labor. Entrepreneurs are constantly engaged in substituting the agents of production for each other: putting more land with the same amount of labor or putting more labor with the same amount of land; putting more capital with the same amount of labor or putting more labor with the same amount of capital. And any entrepreneur tries to employ labor up to the point where its marginal productivity and the wage he must pay are in agreement.)

Substitution such as we have been considering has certain limitations. The capital equipment in an established factory may not be susceptible of much modification. Machines are constructed to be run by a certain number of men, and the entire layout of the plant may be such that additional laborers can be accommodated only within limits. The same limitation applies to the combination of new machines with old equipment. But a much wider range of substitution is possible in new factories. Here the machinery and the layout of the plant may be designed to economize the use of labor, as compared with old plants, or to econo-

mize the use of capital equipment. If new plants can make better combinations of the factors of production than can be made by adjustments within old plants, then the tendency will be in the direction of the abandonment of old plants.

If only a few entrepreneurs were engaged in the experiment of finding out the best combination of the factors, and if experimentation and substitution were unknown throughout the rest of industry, then we should be forced to admit that in these few plants the going rate of wages determined how many laborers it would be profitable to employ with a given quantity of land and capital. (And since the marginal productivity of labor depends, its efficiency being given, on the number of units employed, it would be necessary to admit that the rate of wages determined the marginal productivity of labor.) And this is substantially true for any single plant. The employer cannot appreciably affect the rate of wages by discharging men or by hiring more. The number he will combine with his capital depends in part on the rate of wages; but, for industry at large, wages do not determine marginal productivity. Laborers must work to live, and, barring unemployment caused by maladjustments, sickness, and strikes, the entire labor force tends to be taken into industry. The total supply does not depend on the wages paid (at least not directly), and so the marginal productivity of labor does not depend on the wages paid.

To prove this statement let us suppose that entrepreneurs generally find that unit costs can be reduced by substituting machines for labor. These machines, it must be assumed, are known and in use in this or other industries. These are not recent inventions. Many laborers will find themselves without jobs. Wages will then tend to fall until all capable, willing workers are employed. An equilibrium will be reached when wages have been forced down to the point where labor's marginal product is equal to its wage. One may ask, however, whether it would be possible to substitute capital for labor even after this point had been reached. Might not a machine costing \$100 turn out work worth \$110 while labor costing \$100 turned out work worth only \$100? If this were the fact, the makers of the machine would quickly discover it and raise the price of the machines to \$110. Or if the price could not be increased, more machines would be employed, their marginal productivity would fall, laborers would be discharged, and wages would fall until a machine costing \$100 turned out a \$100 product while labor costing, say, \$95 turned out a \$95 product. It is evident, of course, that some labor would be required to produce the machine.

THE SUPPLY OF LABOR

The number, or the supply, of laborers is one of the conditions determining labor's marginal productivity; the others are (1) the amounts of the other factors available for use with labor, and (2) the ability and capacity of the laborers. Here we are concerned only with the supply of labor.

The supply of labor, like that of every other factor of production except land, responds to the demand for it. If the demand for machines increases, more will be produced; if the demand for savings (new capital) grows, the amount of savings will be augmented; so, too, with labor. But the response is not the same as it is in the case of the machines; the machines will be increased until their marginal cost of production equals the gains that marginal entrepreneurs can derive from them. Applying the same reasoning to the supply of labor, we should conclude that any increase in wages would result in a further growth of population to the point where wages just covered the cost of rearing and training new laborers. (This was essentially the position taken by Malthus and his early followers, who proposed the subsistence theory of wages.)

Stripped of its historical setting and of its connotations in relation to problems other than wages, the argument of Malthus ran as follows: the biological possibilities of growth in population are such that in a very short time any given population is capable of doubling its numbers. In new countries, where this maximum is most nearly approximated, populations have been known to increase 100 per cent in twenty-five years. Because of the sex instinct, the population *tends* to increase at this rate. That it does not always do so is due partly to certain noneconomic checks, such as war and disease; but the principal reason, according to the Malthusian theory of population, is the lack of subsistence—that is, real wages. When wages increase we may expect population also to increase until wages fall again to the point where they are just sufficient to maintain a stationary number of people.

If the argument had stopped at this point, it would have stated a cost-of-production theory of wages. It was noted, however, that among certain classes the desire of parents to give their children a good start in life and their fear of being unable to maintain their own standard of living sometimes limited the size of families. Yet in spite of this exception the reasoning led to the conclusion that the cost of subsistence governed the supply of labor and therefore wages. Since the time of

Malthus the theory of population has been modified, chiefly by greater emphasis on the voluntary control of the birth rate and on the effect of the noneconomic factors that determine it. Among the nations of the Far East it is perhaps true, even at the present time, that population keeps close to the maximum permitted by the means of subsistence; on the other hand, among the peoples of western Europe and among their descendants in other parts of the earth other checks than want of subsistence are at work to limit numbers.

First among these checks is the desire of parents to secure for their children a better economic position than they themselves enjoy. To accomplish this objective children must be better cared for in infancy, so that they will enter the strenuous years of life with better physiques; they must be educated, and they must be equipped either with capital or with special training for a trade or a profession. But all these advantages cost money, and if they are to be provided the size of the family must be limited.

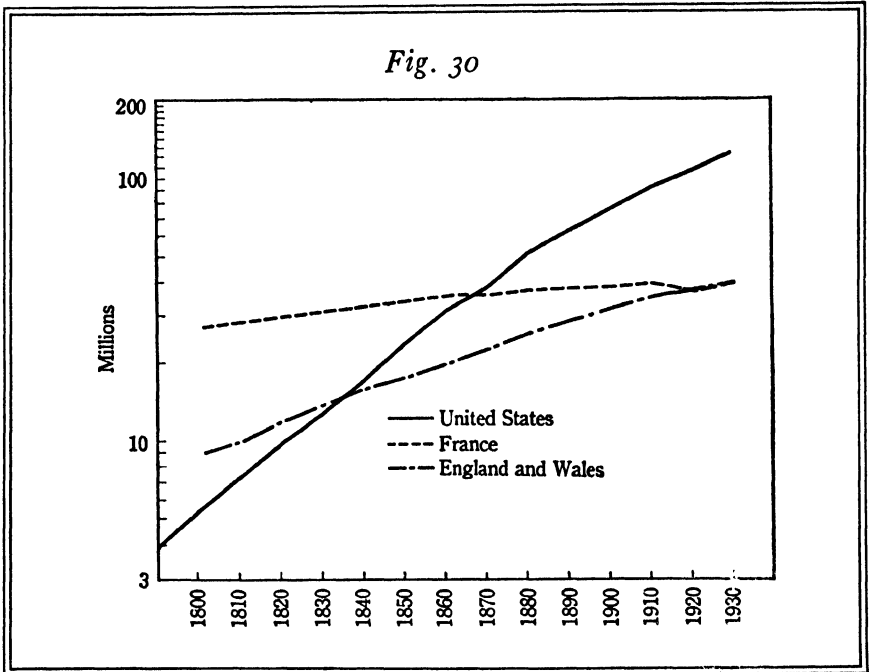
In the second place, change in social ideals and urbanization have made it possible for many women to pursue trades and professions. They are thus able to secure a comfortable income and a satisfactory social position without marrying. The number remaining unmarried through life has not as yet increased to very large proportions; but the age at marriage has increased, and consequently the number of children that can result from each marriage has declined.

A third factor affecting the birth rate is the standard of consumption. Coincident with rising incomes has come the demand for more and more luxuries; and this rise in the level of expenditures, it is supposed, also has caused postponement of marriage and limitation of families.

As yet these conditions have proved most effective in limiting the birth rate among the middle and upper income classes, but they are becoming more and more effective among the lower income groups. Nevertheless there is still much truth in the Malthusian doctrine that the supply of labor varies with the demand for it. It increases when improvements and discoveries make living conditions easier. And one of the results of the industrial progress of the nineteenth century was a great increase in the population of all the countries of western Europe and of the two Americas.

Fig. 30 shows the growth of population in England and Wales, France, and the United States from 1800 to the present time. The use of the logarithmic scale enables us to see at a glance the *rate* at which population accumulated in the three countries. The greatest rapidity

of growth has been shown by the United States, and the least by France. Yet at the beginning of the nineteenth century even France showed the same tendencies toward increasing population as are found in England and the United States. Probably the most important cause of this increase, an increase common to all western Europe as well as to the countries for which the data are given, was the increased demand for labor. Two conditions were at work to increase this demand: In the



first place, the inventions that characterized the Industrial Revolution made each worker's labor much more effective. This led to an increase of goods, easier living conditions for the entire population, and a higher rate of increase in numbers. As might have been inferred from the Malthusian theory, population increased because the means of living increased—not only food, but other necessities of life as well.

The second factor that stimulated rapid additions to the population was the opening of the fertile farm lands of the Americas and the discovery on those continents of rich mineral resources. The co-operation of the railroad and the steamship made these natural resources of the Western Hemisphere available for the use of the people of Europe. When transportation became cheap, first England and to a less extent

France, then Germany and other countries, began to import large quantities of food and raw materials, which the early development of transportation in America delivered cheaply in the Old World; and by cheap transportation the finished commodities made from the raw materials were dispatched to the most distant parts of the globe.

England led in the movement. She became a nation largely specialized in manufacturing, for which she had advantages in the possession of iron and coal, a large merchant marine to bring in raw materials and to send out finished commodities, a long experience in manufacturing, and a large colonial empire to serve both as a market for manufactured goods and as a source of materials. In France specialization was not carried so far. The population tended to remain more agricultural, and in manufacturing small-scale plants devoted to the production of specialties prevailed rather than large mills, ironworks, and shipyards. With her larger area France might have had a much larger population than England, even without England's natural resources, if her people had turned early to mass production; but since the people retained their attachment to agriculture and to the production of the finer sorts of manufactured products, and since their failure to secure an early start put them at a disadvantage in comparison with England, France reached the limit of growth in population before England.

In the United States we see a more rapid increase of population than in England and France. Both the natural increase of numbers and immigration have contributed to this growth. But the reason for both sources of additional numbers is the same: high labor productivity owing to bountiful resources and to the adoption of machine methods and large-scale production. In other new countries, such as Canada, the same rapid growth of population has taken place. Given fertile land, primeval forests, untapped mineral resources, and a healthful climate, population growth waits only the establishment of government and the development of transportation.

In the century and a quarter of growth in population shown by the chart we see at least a partial vindication of the Malthusian theory. When the supply of consumable goods became more plentiful, population did increase, as Malthus had predicted. And it increased most in the country where the additional supplies of commodities were most readily available, and in those which were best fitted by institutions and by past experience to reap the gains of the new techniques of production that the Industrial Revolution made possible. In France, where the new techniques were not so readily adopted, the increase of population was less.

DECLINE IN POPULATION GROWTH

Until recently considerable alarm was voiced by many writers on population lest the continued increase in numbers should result in a decline in the standard of living. During the decade after the World War a sharp reversal of opinion occurred, and many competent authorities now fear that the twentieth century may yet witness a considerable decline in the populations of countries in northern and western Europe. This shift of opinion is due to the fact that it is now known that, on the basis of existing fertility and mortality rates, these populations cannot long continue to replace themselves.

At the moment the *age distribution* is favorable to a relatively high crude birth rate. It should be understood, in this connection, that the crude birth rate and the crude death rate are calculated in terms of the number of births and deaths per thousand of population, regardless of the age distribution. There is at the present time in all these countries a disproportionately large number of persons aged 20 to 50. This distribution follows in part from the fact that this segment of the population was born in a period when the populations of these countries were increasing rapidly, and hence persons of these ages are disproportionately numerous relative to those who are older. On the other hand, during the last twenty years there has been a drastic decline in the birth rate, so that the number under 20 years of age is relatively small in comparison. Now, since it is just the persons aged 20 to 50 who rear children, it follows that the current crude birth rates are abnormally high, owing to the current abnormal age distribution. In another generation, when this peculiar age distribution has been eliminated, the birth rate will fall, even though there should occur no change in the meantime in the age-specific fertility rates (that is, in the fertility rate per thousand women of each specific age).

Moreover, since a slowly growing, stationary, or declining population necessarily will have a larger proportion of old people, it follows that the death rate is bound to rise, even though the age-specific mortality rates do not change. Thus in another generation it seems not improbable that the death rate may exceed the birth rate. Should this develop, populations will begin to decline.

The *net reproduction rate* corrects for abnormalities in age distribution. Based on age-specific fertility and mortality rates, it reveals whether or not a population can be expected in the long run to maintain itself. These rates have been calculated for countries of western and

northern Europe, and it appears that, once the abnormal age distribution disappears, these populations will decline from one generation to the next at the ratio of 1.00 to 0.76. This rate "means that according to fertility and mortality for 1933, 100 mothers give birth to only about 76 future mothers."¹ In the United States, it is estimated that the net reproduction rate is slightly less than 1.

Whether or not an actual decline in the population of western and northern Europe will in fact occur within the present century, it is evident that the populations of these countries are increasing less rapidly than those of Russia, Japan, and Italy. The net reproduction rate in Italy, while falling, is still 1.18, and for Russia it is 1.7. Russia is growing at the rate of 4,000,000 per annum, Japan at 900,000, and Italy at 400,000 per annum. It is probable that the rate of population increase in Italy will decline materially in the next decade, but it appears less probable that any substantial decline will occur in the near future in Russia and Japan. While the birth rate is now falling in Japan, the death rate has fallen proportionally. These population trends in various countries, it is not difficult to see, are bound to have an important bearing not only on standards of living but also on future international relations.

THE SUBSISTENCE THEORY OF WAGES

The rigid application of the Malthusian theory of population led to what has been called the subsistence theory of wages. The working classes, it was said by the advocates of this theory, naturally increase to the point where their wages are just sufficient to enable them to live and rear enough children to keep the population constant. If real wages increased, then the numbers of the workers also would increase and wages would decline to the level at which the family income would be only sufficient to maintain the population at a level of consumption which was little above the necessities of life.

This pessimistic theory condemned the majority of the population to a hopeless round of toil and dull surroundings. Its principal error is found in the overemphasis on the tendency of the working classes to destroy for themselves all gains from improvements in technique by increasing the population. Like many early theories of economics, this one failed to take account of the complexity of the forces governing the supplies of the factors of production. Doubtless the population

¹ Robert R. Kuczynski, *Population Movements*, p. 45. (Oxford University Press, 1936.)

can increase rapidly enough to offset any gains in marginal productivity that may come from accumulations of capital, the discovery of new resources, and improvements; but whether it will do so depends on all the forces that govern the growth of numbers. And, as we have seen, the supply of food and other necessities of life is only one of the external conditions that affect additions to the population.

Another error in the subsistence theory was the overemphasis on the practical consequences of the law of diminishing returns. During the first half of the nineteenth century, and even later, economists who adhered to the subsistence theory often neglected the limiting conditions that must be present before diminishing returns result from the application of more and more variables to a fixed factor. With improvements in technique the fixed factor did not remain fixed, and each laborer came to have more resources with which to work, and not less. From Ricardo (1817) onward it was recognized that in manufacturing, increasing rather than decreasing productivity was likely to accompany population growth. But it was insisted that land was limited and that in the production of food the effects of diminishing returns were unescapable. Hence the increasing population would be unable to increase the supply of food except by a more intensive utilization of the limited supply of land. Real wages—that is, the goods and services purchasable with money wages,—inevitably would fall. This reasoning overlooked two important conditions. Land doubtless was limited, but as time went on improvements in transportation made available for the people of Europe the produce of the agricultural areas of other continents. Moreover, the invention of labor-saving machinery for use on farms, the development of more prolific varieties of grains, and the general improvement of methods of agriculture both in Europe and in America greatly increased the available food supply.

Another factor neglected by the subsistence theory was the effect of rising incomes on the standard of living. When the incomes of the laborers increased, the standard of living also increased, and the growth of population was retarded to such an extent that the biological maximum was not even approached. The rising standard of living was accompanied by a falling birth rate, although other social conditions also contributed to this result.

SUMMARY OF THE THEORY OF WAGES

The general level of wages is determined by the demand for labor and the supply of labor. The demand for labor, which means the amount of labor the entrepreneurs stand ready to employ at a given wage, depends on the marginal productivity of labor. The marginal productivity of labor is determined by the supply of labor and by the supplies of the other factors with which it is combined, the state of productive technique being given. The demand for labor therefore depends in last analysis on the supplies of the other factors. The state of productive technique, or, as it is frequently called, "the state of the arts," depends on all the forces that make for the cultural development of a people. These lie, for the most part, beyond the field of economics. The supplies of laborers depend upon the forces that govern the numbers of the population. Some of these forces are strictly economic and others are not, or at least they are not entirely so. The supplies of the other factors of production—land, capital, and entrepreneurship—are determined in part by natural and unalterable facts. This is true of land in the broad meaning of the term, and it is especially true of mineral deposits. But here, again, we must not overlook the effect of improved technique, which cannot increase land area but which often increases the productivity of a given area of land. The supply of capital will be spoken of in a subsequent chapter; it is regulated in part by the necessity of waiting and in part by the facilities for directing savings to productive uses.

The increase in population shown in Fig. 30 was accompanied by greater productivity per worker arising from specialization, inventions, and new resources. In all the countries of western Europe there was a gain in real wages during the nineteenth century. It was greatest in those that were able to make the greatest gains in productivity relative to the growth in population. England and Wales, as we have seen, gained more in population than did France, and yet wages did not fall below the wages of French workmen; in fact, the wage level of the former country probably rose more than that of the latter. This was because the relative increase in productivity was greater in Britain than it was in France. Had the population of France increased during the nineteenth century at the same rate as the population of England and Wales, it would have been impossible to maintain wage levels unless she had adopted specialization in the manufacture of goods produced by large-scale establishments. Even had she done so, she would not have

been able to maintain the same rate of increase in numbers without reducing the standard of living, because she lacked resources equal to England's.

The marginal-productivity theory of wages is not a final explanation of wages: it is an analytical device for explaining how the forces of demand and supply are related in determining wages, *given the supplies of all the factors and the state of productive technique*. It also throws into clear relief the forces that do determine the course of wages through a period of time. Moreover, it must be remembered that it is only a general explanation. It deals with mass data and explains general, or typical, relationships and results. It supplies the analytical tools for the examination of any wage problem.

IMPERFECT COMPETITION AND WAGES

In the explanation of wages given in the preceding portions of this chapter, competition was usually assumed to be perfect. The employers were supposed to sell their products and buy the factors of production under conditions of perfect competition. Under these conditions prices would always equal cost of production, and wages would always be at the levels indicated by the marginal productivity of labor throughout all industry. Moreover, the relative wage rates paid for labor would be the same in all industries, differences in natural ability and cost of training aside.

If the producers of an important commodity possess a considerable degree of monopoly power, that is, if competition in the market for their good is imperfect, there is always a tendency for the demand for the factors of production in that line of production to be less than it would be if perfect competition prevailed. But if these factors have many other uses (have great mobility), their prices are not likely to be affected very much; nor are they likely to be unemployed. The laborers, the capital, and the land that would have been used if competition had been perfect will be distributed among all the other industries in the country. Since the amounts of the factors kept out of use in the monopolized industry are probably unimportant in amount in comparison with the total of these factors in all other industry, the adding of them to all other industry will not have an important effect. Of course, if a large portion of all industry were controlled by effective monopolies, the productivity of the factors would be reduced by a large amount.

By means of trade-unions, laborers can create imperfect competition in the supply of labor in certain types of production or in certain trades that are found in many types of production, as has already been shown. If a trade-union, say the building trades, can limit the supply of workers by the exclusion of men who may wish to become employed in building, they can thereby raise their wages above the level of the wages of workers of equal natural ability and skill in unorganized trades. If we assume that the capital and management engaged in construction are subject to competition from outside, these factors of production will receive the competitive return received in outside employments. But the labor which has now been made scarcer and dearer in this industry by reason of the restriction imposed by the union will receive more than the competitive return. The scarcer a factor is, the more it will be economized and the less of it will there be to combine with the other factors. (From this it follows that the marginal productivity of labor in this restricted industry will be higher than that of labor in industries and trades where no restriction exists.) Wages will be above the competitive level in the industry. Building costs will be higher, and the tenants will be compelled to pay higher rentals.

If the restrictive effects of trade-unions extend to only a few trades, or if the restriction is very small, the number of laborers who might otherwise have entered the restricted trades but who are kept out will be small also. But if the restriction is widespread, or very great in important industries, then the number of excluded laborers will be large. These excluded laborers will find that their opportunities for employment will not be as good as though no restriction existed. The reason for this diminution of opportunity is clear when we recall the fact that in the restricted industries there will be relatively more land, capital, and management per laborer than if perfect competition prevailed. If that is so, then there must be less of these factors per unit of labor in the unrestricted industries, and therefore the marginal productivity of labor in these industries will be lower than if no restriction existed.

It sometimes happens that one employer or a combination of employers possesses a degree of monopoly control over the demand for labor in a given market, although they have no monopoly of the sale of the product. This occurs when one employer uses a type of labor for which there is no demand by another employer in the same region. Of course, this labor can usually find work of some sort, in times of normal prosperity, but laborers often prefer to accept part-time work rather than shift to some other occupation. They can usually find employment

in the same trade in other places if they are informed and if they have the reserve fund to meet the cost of moving; but they may be uninformed of the opportunities elsewhere, or they may prefer, for a great variety of reasons, to remain on part time rather than move.

Let it be assumed that the employer is at first operating his factory at optimum output (at the point of lowest total unit cost) and that all the laborers of the type he employs are at work. If now the employer decides to use his monopoly advantage he will lay off some of his employees. If, as is assumed, they are unorganized, competition presently will enable him to reduce wages. But when output is reduced the overhead cost per unit of product will rise. If it rises more than wages fall, the employer will be forced to increase employment and raise wages. But if the overhead cost rises very little while wages fall a great deal, it obviously will be profitable for the monopolist to reduce employment. How far he will go in the direction of curtailing his demand for labor depends, as has been suggested, on the extent to which wages can be reduced by a relatively small amount of unemployment (created by him) and on the importance of overhead cost. At some point it will be found that the profit resulting from created unemployment is a maximum. At this point the employer will cease to contract output.

There are several limitations on the use of this type of monopoly power. It is sufficient to explain them briefly, because their effects are readily apparent. (1) If the workers have a strong trade-union, they can successfully resist the wage-reducing tactics of the employer. Obviously he cannot gain by *merely* laying off men and paying the old wage. By this procedure he reduces output and raises the average cost of production. (2) If the state regulates wages either directly by wage-fixing or by employing all involuntarily unemployed workers at the customary wage of their trade, the monopoly of the private demand for labor is powerless. (3) The permanence of the monopoly of demand for labor depends on the mobility of labor. Usually there is some mobility, and after a time such monopolies would lose their power were it not for dynamic conditions which may create new monopolies or strengthen old ones.

The position of the employer who has a degree of monopoly of the demand for labor is strengthened if he also sells his products in a market protected from the competition of other sellers. By reducing output he is able not only to reduce the wages of labor but also to raise the price of his product. In Chapter XII it was shown that the seller who possesses a monopoly of supply will reduce output to the point

where the marginal-revenue curve intersects the marginal-cost curve. But in that discussion it was assumed that the position of the marginal-cost curve could not be changed by the seller. But if the seller is also a monopolist of the demand for labor, he can reduce the cost of labor and thus move the average variable-cost and marginal-cost curves downward and to the left. In this manner he can create two sources of monopoly profit, one by raising the price of the product and another by reducing the wages of labor. In this case the effect of a monopoly of the demand for labor is also to cause unemployment and to reduce the national income.

There are a number of other types of imperfect competition that affect employment and wages. Laborers are immobile and therefore the same sort of work in different parts of the country may be paid at different rates. Capital and management also are immobile at times, and this too may affect wages. Some of these effects are discussed in the next chapter.

WAGES AND IMPROVEMENTS IN TECHNIQUE

Improvements in productive technique have usually enabled man to obtain from nature a greater return for his labor. In terms of Fig. 29 such an improvement raises the curve AC throughout its entire length. Hence, with any given supply of labor before the improvement takes place, an invention raises the marginal product of labor, and wages rise. The effect of an improvement must be sharply distinguished from the substitution of labor-saving machinery for labor, which often takes place when wages rise above the marginal productivity of labor in an industry that employs only a limited portion of the national labor force. In the explanation of the principle of substitution it was assumed that the wages of labor rose in a particular industry because the marginal productivity of labor had become greater throughout industry generally, while the return to capital had not. In the industry where this rise is not warranted by the productivity of labor, the proportion between labor and capital is unfavorable to labor. There is too little capital for each laborer to work with; there is too much labor for the capital employed with it. The proportions must be changed if competition prevails.

But when improvements in the use of resources occur, labor acquires more natural resources to work with and usually additional capital to aid it. This necessarily increases the productivity of labor to some extent. For example, when the steam engine was invented it made

possible the substitution of natural resources for human energy. But although this made it impossible for labor to secure employment in tasks where mere brawn was needed, it also created other employments.

That inventions are beneficial to society has seldom been questioned. One sometimes hears the assertion, however, that the substitution of machinery for laborers reduces wages in the industry where the substitution takes place. Whether or not wages will be reduced depends in part upon what we mean by "in the industry." If skilled mechanics have been employed to make a commodity, and if a machine is invented that permits the use of unskilled workers (perhaps women and girls) or that is semiautomatic, then the wages of the workers directly concerned with turning out the product probably will fall. And the skilled mechanics whose trade is taken away may be forced to accept work of a different sort, at which they will earn lower wages. This latter result need not follow. The introduction of the machine may be so gradual that the skilled hand workers have time to learn new trades, or the rate at which they are displaced may be so gradual that the number of withdrawals necessary to maintain wages at the old level amounts only to the number who would naturally disappear because of death and old age.

The fear once expressed by laborers that machines might permanently displace labor, or at least reduce the total demand for labor and thus lower wages throughout all industry, has not been justified by the experience of the past century. The use of machinery is merely a method of employing labor in production. A worker may attack a piece of work in one of two ways: if he employs direct methods, he will use very simple tools; if he employs indirect methods, he will make machines to aid him. Usually, though not always, the indirect method enables the worker to produce more goods than the direct method. If it does not, it never will be employed.

The record of history shows that the introduction of machinery has not decreased the total demand for labor. Against the possible decrease in demand at the point where a machine is substituted for hand labor must be set three new sources of increased demand. We may take as an example an improvement that displaces a simple sort of labor, say a machine for spraying house paint. When the machine is introduced, some painters may be thrown out of work. But the demand for mechanics to make a machine of this sort will be increased; this is the first new source of demand. Second, if many painters are displaced, the cost of spreading paint also is reduced, and, in consequence, there is a

greater volume of demand for painting materials and for operators of the spraying machines. Thus, if the demand for the product in the manufacture of which the improvement occurs is very elastic, then the fall in price resulting from the improvement will lead to increased buying, larger output, and re-employment of the displaced laborers. Third, a great drop in the price of the service, which may follow from the invention, will free a portion of the consumer's income to be expended in the purchase of other goods.

The slope and shape of the demand curve for the product of the industry in which the improvement is made has a great deal to do with its effect on employment and wages. If the demand for the product is very inelastic, and if the improvement considerably reduces the number of laborers required in the industry, this reduction will be permanent. Because of the inelasticity of the demand very little more of the commodity can be sold, even though the price is greatly reduced, and very few of the laborers ousted by the new machine will again find employment in the industry. On the other hand, if the price of the product falls a good deal, and if very little more of it is purchased, consumers will spend a much smaller amount of income in buying it. They will then increase their demands for other goods, and in the industries producing these goods the demand for labor will be increased. But it must not be forgotten that when the workmen are displaced by the machines, they are unemployed and their incomes are cut off. Until they can be reabsorbed into other industries there will be little or no net increase in the demand for goods, and little or no gain to society at large. The displaced workers will lose what the consumers gain. Because there has been no increase in the total demand for goods, the workers displaced by the labor-saving machinery will not readily find new jobs. Yet there is ample historical evidence to show that such workers somehow are reabsorbed. In England, during the nineteenth century, the percentage of unemployment probably remained practically constant in spite of the rapid introduction of machinery. There was some unemployment, of course, but it was the result of maladjustments incident to a dynamic world. Since the percentage of unemployment did not increase, we are forced to conclude that the introduction of machinery did not permanently displace labor. Since the World War, however, the percentage of unemployment in England and recently in the United States has increased.

The persistent unemployment in England since 1920 is largely an unsolved problem. It has been due in part to the inability of the manu-

facturers to reclaim markets lost during the war, to artificial interferences with wage rates, and to the immobility of labor. In the United States the persistence of unemployment in the face of increasing production since 1933 likewise is a troublesome problem. There is reason to believe that when wages are forced above the marginal productivity of labor and when new savings are abundant, labor-saving machines may take the place of labor and thus give rise to unemployment. It is also conceivable that in a society otherwise static, inventions displacing labor may cause unemployment unless wages are reduced. The problem is extremely complex, and as yet there is no really satisfactory statement of the relation of inventions to unemployment and wage rates. In the past century and a half, improvements in technique and rising real wages have come in the same countries and at the same time. But during this period two new continents were mostly settled and many new resources were made available for labor to work with. Hence the historical record does not prove that inventions must always be favorable to rising wages.

The effect of inventions upon the workers in particular industries is much clearer. When the inventions came in the glass industry, workers were displaced. They could not be reabsorbed, despite a relatively high elasticity of demand for the product. Hence they were forced into other industries. This condition is typical of the effects in many industries. It is conceivable that the unemployed workers may by competitive bidding for jobs lower wages in other industries and in that manner create an inducement for employers to combine more labor with the other factors of production. But in modern societies with trade-unions, unemployment insurance, and other institutions that tend to bolster up the resistance of the workers to wage reductions, this type of adjustment cannot always be made. At least it is difficult to bring it about within a short period of time and to a sufficient extent to absorb all the workers displaced by the machines. Under these conditions the reabsorption of the workers must await the development of new industries, new products, and new types of service. But these new sources of employment do not come into existence spontaneously. It requires inventive skill and new capital to develop new products, to stimulate new wants, and to get new industries going. New fields must be explored, consumers must be motivated to desire new products, and those who have funds to invest must be convinced that the new projects are likely to become successful. Advertising plays a part in the adjustment, and so does bank credit, which enables the entrepreneurs to put the

productive forces in motion in anticipation of the product out of which eventually the factors of production must be paid. After the new industries have become productive, the additional output of goods which they turn out increases the income and the real purchasing power of the entire country. Hence it comes about that the labor-saving improvements do finally increase the incomes of the laborers as well as those of other industrial classes. But it must not be forgotten that the workers who were displaced by the machinery could not have contributed to the income of the community until they were placed in a position where they could earn an income through productive labor.

PRINCIPLES OF WAGE ADJUSTMENTS

Frequently wage rates are set by collective bargaining or by boards of arbitration. The factors usually considered in these adjustments are: (1) changes in productivity, (2) changes in the cost of living, (3) the minimum wage, (4) wage differentials, and (5) wage standardization. The last three of these factors are considered in the next chapter. Here we shall deal with the question whether wage adjustments according to the first two factors are or are not in accordance with the theory of marginal productivity.

In arbitration cases it often is argued that an increase in the productivity of an industry is a cause for higher wages only if it is due to an increase in the personal efficiency of the workers; but that if the cause is improved machinery or better management, wages should not be raised. According to the marginal-productivity theory, labor also should participate in the gains arising from increased productivity caused by additional quantities of the other factors or improvement in them. Of course, labor is supposed to benefit only to the extent that it becomes marginally more important. And historical evidence given in Chapter XXIX shows that a given percentage rise in the per capita physical productivity of the United States has been accompanied by a nearly proportional rise in wages. It is not, however, the productivity of a single group of workers that determines their wages, or the wage level, but the physical productivity of industry as a whole.

The "cost of living" principle of wage adjustment means that if no other factors are affecting wages, money wages should be varied as the cost of living varies. The marginal-productivity theory deals with real wages and is in entire accord with this principle, since a failure to adjust wages to a rise in the cost of living would mean that real wages

were reduced. However, the adjustment of money wages to the cost of living merely serves as a starting point from which further adjustments may be made if justified, for example, by an increase in the marginal productivity of labor.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It is sometimes said that the laborers could increase their wages by first increasing their standard of living—that is, by demanding more and better goods for consumption. Is this statement correct?

2. If we define the standard of living as the sum of goods and services that people demand before they will assume the obligations of rearing a family, then, in the long run, the standard of living will affect wages. Show that this statement is in agreement with the marginal-productivity theory of wages.

3. How should you explain the fact that as the population of a new and sparsely settled country becomes denser, manufactures become cheaper? that the average productivity of the population increases up to a certain point?

4. "There is just a certain amount of work to be done—just a limited number of jobs in the world. If one man does more than his share of this work, there will be less for others to do. If he does as little as he must, then there will be more work for the others, and wages will be higher." Show the fallacy of this reasoning.

5. It has been argued at various times that if the rich should save less and consume more, wages would rise. Show by the use of the marginal-productivity theory that this reasoning is fallacious.

6. Does the substitution of capital for labor raise or lower the marginal product of the labor that remains employed? Does it raise or lower the marginal product of the capital?

7. Does the theory of marginal productivity explain the wages of those who render direct personal service, such as singers, physicians, hairdressers, and waiters? Are the wages of such workers affected at all by the quantity of land and capital available in the community?

CHAPTER XXIV · Wage Differences



The preceding chapter dealt with the general relations between the supply of labor and the other factors of production. It was concerned with the most general conditions affecting the normal rate of wages. It did not take account, except incidentally, of the fact that the earnings of labor vary a great deal from occupation to occupation and, to a less extent, in the same occupation from one geographical area to another. The present chapter, which carries further the discussion of imperfect competition in distribution, is concerned with some of the reasons for these variations and with the relations between the supply of labor and the other factors within different occupational groups.

LABOR-INCOME GROUPS

It is a matter of common observation that there exist at the present time several fairly distinct groups of workers. The lowest-income group includes all those who perform common labor. This group is characterized by lack of special skill, usually by a scanty education, and often by low natural ability. The male workers of this group do most of the heavy manual labor about factories and stores and in the construction of buildings, railroads, and highways. They supply the bulk of casual labor for all outdoor occupations. The women workers of this group are employed at simple manual tasks, and as saleswomen in the cheaper shops. Both the men and the women of this group are employed to tend automatic machines and to operate simple mechanical devices.

Not very much above the common laborers are the semiskilled. Among the male workers of this class a somewhat higher degree of specialized skill is required, or at least a familiarity with somewhat more complicated machinery; in clerical work and in salesmanship they occupy higher positions or work in better establishments. For most of these tasks either greater native ability to learn or a longer period of training is required than for common labor. Evidently the term "semi-skilled" does not apply strictly to clerical work and selling; but the degree of attention, of independent self-direction, and of training among

semiskilled workers in mechanical trades and in the other two lines of work is roughly comparable.

Above the semiskilled workers are the skilled mechanical trades and the upper groups of clerical workers and salesmen. More training and still greater native ability is required for these occupations than for common labor. The worker must have more ability to think for himself, and he must accept more responsibility for his work. Among the male workers in this group we find the skilled craftsmen in construction and in the machine trades: carpenters, masons, electricians, lathe-operators, diemakers, and the like. In clerical pursuits are bookkeepers, secretaries, statistical clerks, and stenographers. The skilled workers also include salesmen and the engineers, firemen, and conductors on railroads. The skilled group is far less homogeneous than either the common-labor group or the semiskilled group, and the variation in wages is far greater.

The highest-income group is made up of professional people and business executives. Here the period of training is much longer than for other groups, and the average native ability probably much greater. Here are included engineers, chemists, accountants, lawyers, physicians, heads of government departments, and the managers of business concerns. But to a greater extent even than the skilled workers this group is lacking in homogeneity. Some of the practitioners of the professions are no better paid than highly skilled mechanics, and many of them receive less than the most successful salesmen. It is among the managerial class, however, that lack of homogeneity is most striking. Many managers of minor departments in large enterprises and many self-employed managers, such as small storekeepers, farmers, the owner-operators of garages, and small contractors, do not receive larger incomes than skilled mechanics. In a majority of these cases either the amount of managerial capacity required is very small or the worker is an unsuccessful self-employed manager. The small storekeeper, for example, must spend most of his time and energy in doing about the same work as an ordinary salesman and clerk, and the typical farmer is engaged chiefly in ordinary manual labor.

Although this group overlaps a great deal in its lower ranges with the next lower one, it rises far above it. Each of the four groups contains many workers who could do as well for the present in occupations of the next lower group. Some of these are young people who in reality are apprentices learning the tasks of the higher groups, and it makes little difference in which income group they are classified. A junior

accountant may receive a salary that falls below the wage of a machinist, but he will receive higher wages if he makes a success in the higher group.

In considering the causes of the differences in income of these groups, three very important facts that partly offset the observed differences must be taken into account. The age at which the worker can command the average wage is higher in each succeeding group. The common laborer attains his maximum wage almost as soon as he has reached physical maturity; the semiskilled worker somewhat later, though not very much later. The skilled worker requires several years of training. Among the professions from three to six years of college education is common. A business executive often serves an apprenticeship through many years, climbing from minor positions to those requiring more responsibility and business judgment. On the other hand, the common laborer is likely to suffer a decline in his earning capacity at an earlier age than the skilled or professional worker. But the fact still remains that the laborer can command his maximum income at an earlier time; and people generally prefer a present to a future income of the same size, and a small income which can be obtained at once to a somewhat larger one for which they must wait.

The second fact that tends to reduce the apparent differences in wages is the money expense and the real cost of gaining a place among the better-paid groups. The boy who is apprenticed to learn a skilled trade usually must have some additional means of support during the learning period. The practitioner of a profession must incur expenses for his formal education. In addition to the money expense of training, most of the better-paid occupations require greater concentration and more attention to the work in hand than the tasks of common labor. Hence the real cost of learning them is greater.

In the third place, the average person believes that he is much more certain to succeed in an occupation which is followed by many people and in which the pay is relatively low than in a higher occupation. The risk of failure is much greater for the average man or woman, regardless of formal training, the higher the wages paid by the work to be done. And this risk deters those who prefer a smaller but certain income from attempting to reach higher groups, in spite of the fact that they may have the native capacity to succeed in them.

In the fourth place, many occupations have advantages other than the money wage the worker receives. The agreeableness or disagreeableness of the task is a factor in the determination of the supply of labor

for any given occupation. The agreeable types tend to become overcrowded, and the wages of workers in them fall below the level of the wages of the disagreeable. It must be noted, however, that when disagreeableness is coupled with low requirements as to native ability and training, wages may still be low. None of these conditions that we are considering operates independently, and none is a sole determinant of differences in wages. But after all allowances have been made for conditions that tend to offset differences in money wages, there still remains a great deal of actual variation in the labor incomes of different groups of laborers. It is to these differences that we shall now turn our attention.

THE PROXIMATE CAUSES OF DIFFERENCES IN WAGES

Differences in labor incomes, which we have been considering, arise under competitive conditions. They are due to relative differences in the demand for and the supply of labor for the different kinds of work. The price paid for human services, like any other price, depends on demand and supply. (The wages of common labor are low because, relative to the demand for such work, the supply is large,) and because the supply is large the marginal productivity of this kind of labor is low. Capital and land are but passive instruments of production, and the former in its primary form of purchasing power has a high degree of adaptability. Land also can be adapted to different uses. Neither will be used in any industry unless it yields as great a return as could be obtained from it in alternative employments. So we find that the capital and the land used by common labor must each yield the same return as that used by highly paid labor. But since common labor is relatively plentiful, the prices of the products to which it contributes its productive energy will be low in comparison with those that are the products of skilled labor; or if the products are the resultants of both common and skilled labor, the price that must be paid for the former will be low in comparison with the price paid for the latter.

Where the different types of workers are all required to bring a product to the consumers' market, common labor becomes one of many agents in production. Skilled labor, professional work, and managerial work, as well as capital and land, are the other agents with which common labor is combined. Since this grade of labor is plentiful in comparison with the other agents, its marginal product is low, and its earnings are low in comparison with the earnings of the other labor groups, who are less numerous relative to the demand for their services.

The salaries of professional men and managers are sometimes said to be high because these workers accept heavy responsibilities. It is said that the president of a large corporation receives a large salary because his work is very important, because his decisions involve large amounts of money. But responsibilities appear important to us when there are few who can successfully assume them. A pilot who guides a modern steamship into port is responsible for a very valuable piece of property and for the lives of many people; yet his wage is not high in comparison with the salaries of those who hold managerial positions in the steamship company; and the principal reason is that there are many people who readily learn the pilot's trade, but few who can or do qualify for the managerial positions.

Differences in wages are said to be due also to differences in the efficiency of workers. This is a correct explanation of variations within the same trade or income group, but it does not explain the variations between groups. If all common laborers, by an act of magic, were transformed overnight into supermen who could do four times as much work of the same sort as they now perform, their wages would be lower per unit of energy put forth. Each unit of work would have lower marginal productivity, the supplies of the other agents of production remaining the same as they now are. This follows from the law of diminishing returns. Individual earnings per worker per year might be increased after industry had been adjusted to the change; but whether these earnings remained higher would depend upon the maintenance of a relatively small supply of common labor in comparison with the other agents of production.

FUNDAMENTAL CAUSES OF DIFFERENCES IN WAGES

An explanation of the differences in labor incomes among different groups of workers must discover the forces that regulate the relative supplies of and demand for workers of the different classes. The emphasis must be placed on the relative numbers in any group and not on the absolute numbers. In the first place, the labor of all grades engaged in producing a given product may be plentiful or scarce relative to the demand for the product, and in consequence its marginal productivity may be either low or high; in the second place, the laborers of a given occupational level, say common labor, may be numerous or few relative to the other laborers of different groups and the other factors whose co-operation is necessary for the production of a given

commodity. A good may have a high price, and yet the wages of the common labor engaged in making it may be low because such labor is plentiful relative to the skilled labor required for making the good. As has frequently been observed, different grades of labor do not constitute one common fund from which may be drawn at any time any number one pleases of laborers of different sorts. Each group is a partly independent agent of production whose supply is governed by somewhat different conditions. Hence common labor is combined by the entrepreneur with skilled and managerial labor in the same way that labor in general is combined with capital and land. And if the supply of common labor in any industry is plentiful while the supply of the other groups is scarce, the marginal productivity of the former will be low, and its wages will be low in comparison with the latter.

One of the fundamental causes of disparity in the labor earnings of individuals and of groups is the differences between individuals in native ability. Even under uniform external conditions there are only a small number of people who are able to excel in any kind of competition, a much larger number who can attain to moderate proficiency, and a small number who can never get beyond the poorest sort of performance. Hence we should expect that where competition is present only a few people will be found at the lower levels of any sort of work, and that only a few can qualify for the more difficult tasks. But, as has frequently been observed, the distribution of human ability is not closely comparable with the distribution of labor incomes. The high salaries paid to executives and the high fees of some professional men appear to be out of proportion to any difference of ability that can be tested objectively.

There are several conditions that account for the disproportionate numbers of the low-wage groups. An individual can attain to the higher groups if he is equipped with more than ordinary native ability, if he has been carefully trained, and if his opportunities for learning the better-paying sorts of work are good. These three—native ability, training, and opportunity—may be combined in almost any proportions.

Training and opportunity give the children of the skilled artisan better chances of becoming skilled artisans, of entering the professions, or of obtaining managerial positions in business than the children of common laborers. The children of the professional and managerial group have still better chances. And while it may be generally correct to say that true genius is difficult to stifle, it is not correct to say that the height in the economic scale to which an individual will rise during

his lifetime is not affected by the conditions under which he was reared, the extent of his education, and the opportunities open to him to learn something more than a routine, manual sort of work.

There is a tendency for the different income groups to be self-perpetuating. The children of laborers tend to remain laborers, because they inherit a relatively low capacity for anything but simple tasks, because their early training and formal education have not been such as to fit them for anything but that sort of work, and because they have fewer opportunities to come into contact with those in business and the professions who might take an interest in them and provide them with opportunities for the better-paying positions. Finally, they are not brought up in an environment where success is both expected and attained. On the other hand, the children of the higher-income groups doubtless have somewhat better inherited capacity on the average, they are better cared for in infancy and so have stronger and healthier bodies, the period of their formal education is generally longer and better adapted to fitting them to hold good positions, and their opportunities for finding openings that lead to these positions are better. They are also taught from childhood that they must succeed in the world.

The population of any country thus tends to be divided into non-competing groups, or into economic classes. The barriers between these classes are not impassable, however, and certain institutions and economic conditions help the individual to overcome them. First in importance among these institutions are the publicly supported school system and other publicly provided educational facilities. Public supervision of health and public control of sanitation also are factors. Still another factor is the increase in the general level of incomes. If the annual income of a nation should increase absolutely, the incomes of common and semiskilled laborers also would increase, and some, at least, of the disadvantages of the children of the laborer's family would disappear.

DIFFERENCES IN WAGES BETWEEN DIFFERENT PLACES

Everyone is aware of the fact that the same sort of work is not paid at the same rate in different parts of the country. Very often, however, these differences are more apparent than real, even when money wages alone are considered. A trade may go by the same name, and to the casual observer the workmen who follow it in different places may seem

to be doing exactly the same sort of work; but on closer investigation one often finds that there are variations in the skill required or in the conditions of work that account in part, at least, for variations in the rate of pay. And if we look about us we shall find the same variations within a relatively small area. On the other hand, there are many actual variations.

TABLE 25. Hourly Wage Rates of Building-Trades Workers in Eight Cities in 1928 and 1933¹

CITIES	CARPENTERS		ELECTRICIANS ²		STONEMASONS		PAINTERS		PLASTERERS	
	1928	1933	1928	1933	1928	1933	1928	1933	1928	1933
Atlanta . .	\$0.80	\$0.90	\$1.00	\$1.12	\$1.40	\$1.12	\$0.85	\$0.85	\$1.25	\$1.00
Birmingham	1.00	.75	1.25	1.00 ³	1.50	1.00	1.00	.75	1.25	1.00
Boston . . .	1.25	1.18	1.25	1.25	1.40	1.30	1.25	1.12	1.50	1.38
New York . .	1.50	1.40	1.50	1.40	1.75	1.65	1.50	— ⁴	1.75	1.50
Chicago . . .	1.50	1.31	1.62	1.50	1.62	1.38	1.62	1.41	1.62	1.38
St. Louis . .	1.50	1.25	1.50	1.68	1.50	1.25	1.44	1.25	1.75	1.50
Denver . . .	1.25	1.09	1.38	.90	1.50	—	1.13	.75	1.50	1.00
San Francisco	1.12	.90	1.12	1.00	—	1.38	1.12	—	1.50	1.25

In Table 25 are shown the wage rates for five skilled occupations in the building trades. While the requirements for skill and craft knowledge doubtless vary from locality to locality, these crafts—four of them very old—where the machine has done little to aid the worker are largely uniform. The eight cities for which the wage scales are given represent the Northeastern, Southeastern, Middle West, Mountain, and Pacific-coast sections of the country. The table shows that the wages are lowest in the Southeast and highest in the Northeast and Middle West. The other two cities, Denver and San Francisco, occupy a middle position. In part this difference may be accounted for by differences in climate. Construction is seriously interfered with in the North during the winter months, and only slightly restricted in the South and on the Pacific coast. But climate cannot account for the differences between St. Louis and Denver, since weather conditions are if anything in favor of the former, in which higher wage scales prevail. A complete explanation of the variations is not available. In part they

¹United States Department of Labor, Bureau of Labor Statistics, Bulletin No. 604 (1934), *History of Wages in the United States, from Colonial Times to 1928*, Table B.

²Inside wiremen.

³Also \$1.25 part of the year.

⁴Two organizations reported a rate of \$1.25 and two \$1.40.

are due to a generally higher level of wages in the two sections where the wage scale for construction work is highest; in part they arise from differences in the cost of living; and in part from the greater strength of the trade-unions in those two sections.

TABLE 26. Hourly Earnings of Workers in Certain Occupations in the Cotton-Textile Industry, 1928 and 1932¹

STATES	1928						
	DOFFERS (MALES)	LOOM- FIXERS (MALES)	SPEEDER- TENDERS (MALES)	SPEEDER- TENDERS (FEMALES)	SPINNERS (FEMALES)	WEAVERS (MALES)	WEAVERS (FEMALES)
Maine	\$0.352	\$0.548	\$0.390	\$0.390	\$0.311	\$0.424	\$0.397
Massachusetts	.395	.592	.451	.467	.350	.431	.405
New Hampshire	.406	.647	.487	.481	.407	.501	.493
Rhode Island .	.409	.615	.488	.488	.367	.498	.469
Alabama264	.395	.276	.276	.215	.311	.299
Georgia282	.379	.307	.307	.222	.309	.292
North Carolina	.289	.413	.333	.333	.242	.370	.333
South Carolina	.270	.377	.296	.297	.215	.313	.277
<hr/>							
	1932						
Maine280	.463	.296	—	.261	.353	.321
Massachusetts	.347	.530	.381	.286	.289	.366	.336
New Hampshire	.329	.468	.344	.335	.290	.371	.354
Rhode Island .	.286	.423	.274	.332	.238	.313	.310
Alabama224	.336	.231	.215	.181	.275	.263
Georgia235	.338	.257	.254	.195	.280	.277
North Carolina	.245	.377	.273	.263	.194	.298	.276
South Carolina	.220	.327	.245	.230	.166	.272	.262

If we turn from crafts to factory labor the same differences between sections are found. In Table 26 are given the hourly earnings (called "rates of pay" by the document from which they are taken) of five positions in the cotton-textile industry. For two of them, speeder tenders and weavers, the earnings are given for both males and females. In all five occupations, and for both men and women, the earnings are higher in New England than they are in the South. This was true for the years before the depression and for the year 1932, which represented the lowest point of the depression. In 1924, for which no data are given in the table, the difference was still greater than in 1928.

¹United States Department of Labor, Bureau of Labor Statistics, Bulletin No. 604 (1934), *History of Wages in the United States, from Colonial Times to 1928*, Table L (various subdivisions).

Fragmentary data indicate that since 1932 the wages of Southern textile workers have risen relatively in comparison with those of workers in the North. Whether substantial equality between the two regions will eventually come about remains to be seen.

The cotton-manufacturing industry is one that extends along practically the entire Atlantic seaboard. It is a highly standardized industry as far as labor and tasks are concerned. It is also a large-scale industry. Here, if anywhere, we should expect to encounter uniformity of wages. It is evident, however, that there were differences in 1932 in money earnings per hour as high as 50 per cent between the high-wage and the low-wage states. These differences in money wages are offset in part by differences in the cost of living. The workers in New England pay higher prices for housing and some other necessities than those of the South, but real wages per hour are higher in the North.

How are these differences to be explained? They can be accounted for in part by the absence of competing industries in the South. There is less capital to employ all workers, and the productivity of laborers in industry is lower. Until recently most of the finer grades of cloth, which require more skill of the worker, were made in the North. The greater strictness of labor laws and the firmer hold of trade-unions in the North also are factors.

These differences in real wages would be greatly reduced if there were complete mobility of labor between the two sections. But obstacles to freedom of movement exist. The money cost of moving is a factor that keeps laborers in a low-wage district when they know that the rate of pay is higher elsewhere. Often they do not know definitely how great the differences really are. Ignorance of the labor market prevents them from selling their services in the highest market. They do not like to give up associations already formed and to risk not finding conditions any better in another locality. These grounds of explanation apply to other lines of manufacture as well as to the cotton industry.

Sometimes differences in wages between localities are due to differences in the productive abilities of the workers. To an unmeasurable extent such differences in productivity account for variations in all industries.

Since the wares of manufacture tend to be sold in the same markets, it is natural to suppose that the employers of poorly paid labor make unusual profits, that the labor is exploited. Often this is the case, but sometimes it is not. The concern paying low wages may be unfavorably located. Or the management may be inefficient; that is, antiquated

machinery or methods may keep the marginal product of labor below the level of other districts, so that the employer does not receive an unusually large profit. Variations in marginal productivity arising from variations in the efficiency of the management or from the use of obsolete equipment cannot persist unless there is also present some limitation on the freedom of movement of workers. Where limitations are not present, the dwindling of the labor supply for inefficient firms will force the wages paid by them to the same level as those paid by concerns of typical efficiency. The owners of the less efficient plants will lose money; they will not be able to cover all supplementary costs. This is as it should be. Society cannot afford to reward equally the results of good and poor management.

When differences in wages are due to an oversupply of labor in some localities, society will gain if the laborers are redistributed; that is, it will do so if the cost of shifting laborers from one district to another does not exceed the gain in production resulting from such redistribution. When differences are due to inefficiency of management, society will gain if the inefficient managers are forced to surrender control of the factors of production.

THE WAGES OF WOMEN

It is notorious that women are more frequently paid extremely low wages than men and that they often appear to be paid lower wages than men for the same work. Some part of this difference is due to the greater physical strength and endurance of men. Another cause for the lower wages of women is the fact that most of them are not permanent workers. The younger group of women, who do not remain long in industry, must go into work that can be learned quickly and easily. They take up the so-called "blind-alley" jobs. Young men, because they expect to be permanent wage-earners, are more likely to avoid such occupations.

But after all allowances have been made for such differences, the fact still remains that women usually receive lower wages for the same work than men. This difference is chiefly due to the limited opportunities open to women. Workers compare the wages offered in one kind of work with those paid in other kinds of work. They will not accept a lower wage than that offered by another employment. For capable, well-trained men these alternative opportunities are more numerous than for women. In the long run, if women can perform a given

task as well as men, only women will be employed to perform it, and the wage will be lower than if men were employed. From this difference in wages the employer derives no benefit. Competition causes the prices of products to fall, and the consumers who buy the goods receive the gain.

A final reason for the lower wages of women is found in their weaker bargaining power. Since they are often temporary workers, and since they usually have less responsibility for dependents, they are not so easily organized into trade-unions; they do not maintain such strong unions as men and do not fight so tenaciously for high wages.

MINIMUM-WAGE LEGISLATION

The problem of the exploited worker and of the worker whose efficiency is poor because his wages are low has been attacked by means of minimum-wage laws. It is not the purpose of this discussion to describe the provisions of these laws. The principles according to which they operate do, however, fall within the scope of the present chapter. Two dissimilar conditions are encountered in the application of minimum-wage regulation: (1) wages in the industry to be regulated are below the wages paid for workers of equal skill and capacity in other industries; (2) wages are not below those of other industries.

If wages are lower than in other industries, the employers may be taking the difference in profits, and the application of wage regulations will simply force the employers to turn over to the laborers their full marginal product. But often wages are unfairly low, and yet the employers receive no benefits from the low labor cost. Competition among them may cause all the gains to be passed on to the consumers of the products made by the underpaid workers. Here the effect of regulation is to raise prices and drive some employers and some employees into other industries. From this transfer some social gain will result. Before the regulation took effect society was receiving too much of the products of the sort upon which the underpaid employees were engaged, and too little of some other products. The regulation tends to remedy this maladjustment of production.

Underpayment of labor may exist without any gain either to employers or to consumers. The employers may be inefficient as a group, or there may be failure to eliminate the unfit through competition. The inefficient employers cut costs by reducing wages instead of increasing efficiency. Labor, being immobile and redundant in supply, cannot

help itself. Regulation in this case should drive out the inefficient managers or cause them to find another means than wage reduction for cutting down their expenses.

If management is efficient, if there is no exploitation of labor, and if the industry is not overexpanded, the attempt to raise wages by law can result only in driving both labor and capital out of the industry. One effect of the raising of wages by law is to cause employers to select workers with greater care. The least efficient are dropped from pay rolls; the more capable are retained. Therefore the excluded workers will not be benefited; they may even suffer from the operation of the law. Those who are retained will gain, and the consumers of their products will lose.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. What conditions are principally responsible for the immobility of labor? Is mobility desirable?
2. Immigration into a country has the general effect of reducing wages. But a heavy immigration within a short period of time might raise the wages of *some classes of labor*. Which classes?
3. It is a common observation that wages tend to be higher in new countries than in old. How can you account for this tendency?
4. Even if we could provide equal opportunities for all workers,—equal education and equal opportunities to enter trades and professions,—there would still be differences in wages, because of differences in inherited ability. Does it follow from this statement that society should make no attempt to provide equality of opportunity? Would the national income be increased by greater equality?
5. Is it desirable that men should earn no more than women in the same occupations?
6. In old countries where social stratification is well defined and relatively permanent, differences in wages tend to be more permanent. Why?
7. Enumerate the principal forces that tend to eliminate differences between the wages of different groups of workers.
8. If wages for all types of workers were fixed by law at a level above the competitive marginal productivity of these workers, could the law be enforced? How would liberal provision of relief or government work affect your answer?
9. A minimum-wage law might attempt to (a) standardize wages for the same grade of labor within the same industry, (b) standardize wages for the same grade of labor in all industries, (c) raise wages for all grades of labor within a given industry. Which attempt is most likely to succeed?

CHAPTER XXV · The Rent of Land



The two primary factors of production are land and labor. In the preceding chapters we have studied the conditions that determine the wages of labor in general and of different groups of labor. We have seen how labor's share in the common product of all the factors is determined. In this chapter we shall apply the same type of analysis to the share received by the owners of land.

LAND AS A FACTOR IN PRODUCTION

In economics the term "land" includes all the different natural resources that man calls to his aid in production: farms, building sites, waterfalls, coal fields, oil pools, quarries, and mineral deposits. These natural resources differ among themselves in many ways, and yet all have certain physical and economic characteristics in common: all are limited in amount, all are spatially immobile, all are nonreproducible. It is because of these essential similarities that we shall consider them all together and call the share that accrues to them in competitive distribution *rent*.

The similarities of natural resources must not be allowed to obscure their differences. Mineral deposits are nonreproducible, just as farm land is. Unlike the latter, however, mines, quarries, and oil pools can easily be exhausted. Moreover, there is a great deal of variation among the different items of any kind of natural resource in respect to richness and accessibility. Some farming areas are extremely fertile and others are relatively barren; some mines yield nearly pure metal and others only low-grade ores. Some agricultural areas and mineral deposits are near the ocean or other bodies of navigable water; others are buried in the inaccessible interiors of the larger continents. Some are found in those sections of the tropics where an unhealthy or enervating climate makes utilization by civilized man next to impossible. Because of these conditions, naturally rich portions of the surface of the earth sometimes are sparsely populated and little used, whereas areas far less fertile swarm with workers. In some instances the progress of technique has

reversed the relative desirability of two areas. The first settlers in the United States could not transport goods between the Mississippi Valley and the Old World. For them the sandy Atlantic coastal plain and the rocky hills of New England were better land than the fertile fields of the interior. Later the building of canals and railroads reduced the difficulties of transportation; today we regard the Mississippi Valley as the most valuable agricultural area on the continent, and the lands of New England and most of the Atlantic coastal plain are regarded as distinctly inferior in productivity.

THE MEANING OF RENT

In economics the term "rent" formerly meant a payment for the use of land as it exists or originally existed in nature. The improvements that man makes on the land were thought to have their own peculiar return. Yet some of these improvements are, relative to the life of any individual, so durable that they tend to merge with the land itself. Of such a nature are the construction of "made lands" along water fronts, the removal of stones and forests from farming areas, the drainage of swamps, and the leveling of sites for buildings. Hence in more recent writings on economics we find that rent is defined as the payment for the use of land as it originally existed, together with the payment for such man-made improvements as are, for all practical purposes, as permanent as the land itself.

In business usage there is no clear distinction between rent and the return for capital improvements on the land. It is said that a building "rents" for so many dollars a year; the rent is the payment the tenant makes for the use of both the land on which the building stands and the building itself. This return is a payment for (1) the rent of the land, (2) the use of the building and other improvements, such as streets and sidewalks, and (3) the time of the landlord or his agent in supervising the total investment in both land and building. In economic terminology we should say that the payment made by the tenant consisted of (1) rent, (2) interest and depreciation, (3) wages, and possibly (4) profits; and if we wished to speak of the total payment we should call it the "business rental" (or simply the "rental") of the building.

If each landowner used all the land he owned, there would be no payments of rent. Yet the land would be worth just as much in production as though it were all leased to other people by the owners. Rent which

landlords derive from the use of their own land is called *implicit* rent; that which they receive from tenants is called *explicit* rent.

The distributive share with which we are dealing is the *normal competitive* return to land. But competition does not work with faultless precision in the market for land leases. Sometimes the advantage is with the tenant; sometimes it is with the landlord. Hence the rent called for in a leasing contract may be either more or less than the economic rent of land. Some land, especially in the cities, is leased for very long periods of time. When the future rental is agreed upon in advance, there is almost certain to be error in calculating the economic rent, and therefore contract rent does not always agree with economic rent. When we speak of rent, economic rent must be understood.

This discussion of terminology cannot omit one further distinction. Many writers speak of the "rent" of capital goods. Now it is true, as will be pointed out in the chapter on interest, that durable capital goods yield a return that is in some respects similar to land rent; but the identity is never complete, except in the case of those permanent improvements on land which are nearly as durable as the land itself, and hence it is well to limit the term "rent" to payments made for the use of land or to these and implicit land rent. When we wish to imply that the return to capital goods is like the return to land, we shall employ the term "capital rent."

METHOD OF DETERMINING RENT

The explanation of the general theory of wages began with the assumption that a variable amount of labor was added to a fixed amount of land, capital, and entrepreneurship. The marginal productivity of labor was found to be the product resulting from the application of the last unit of the labor force of a country. In the further discussion of the theory it was found, however, that the wages of labor are determined in a competitive, frictionless society through the bidding of entrepreneurs for the use of labor. How this bidding results in a wage level that is the same as that demonstrated by the marginal-productivity theory was explained under the heading The Principle of Substitution (see page 388). The procedure used in explaining wages might be used also in explaining the determination of rent. In the following sections the order of these two approaches to the theory of land rent is reversed because it is more in accord with everyday observation to regard land as the fixed factor and labor and capital as the variables. We shall

begin the explanation with the assumption that land is a fixed factor which entrepreneurs use with variable amounts of the other factors. This approach is sometimes called the Ricardian theory of rent, because it is substantially the method employed by the economist David Ricardo in his *Principles of Political Economy* over a century ago.

THE DEMAND FOR LAND

The demand for land, like the demand for labor, is usually a joint demand. When land is used directly by the consumer for residence, some capital and labor are always necessary for co-operation with it. In the case of producers' use it is obvious that land is never used alone. In our study of value we found that the price the entrepreneur will offer for a good used jointly with other goods depends on (1) the value of the finished product, (2) the scarcity of the good itself, and (3) the scarcity of the co-operating factors. This principle applies to land. A miner working a bed of minerals will offer a high price for the privilege (have a high demand price for the land) if the product is valuable, if the deposits of such minerals are scarce, and if the labor and capital used in mining are relatively plentiful and therefore cheap; if, on the contrary, the ore is low in price, if the deposits are plentiful, and if labor and capital are expensive, he will not be able to pay the landlord very much for the use of the mine.

The demand of all society for available land is affected by the same conditions. The rent of land in a densely populated, industrialized community is generally high, because many workers are seeking to use it, because the population requires a large amount of the products of land for subsistence and raw materials, and because there is much capital as well as many laborers available for combination with the land.

The payment of rent by an entrepreneur depends upon the conditions of joint demand laid down in the preceding section. We may first ask the question, Under what conditions would the entrepreneur pay no rent? It is clear that the mere desire of the landowner to receive rent has nothing to do with the answer. In the first place, the entrepreneur cannot be compelled to pay for the use of the land if the supply of it is so plentiful that the refusal of any particular landowner to allow his land to be used will not create a scarcity of land. So long as land of the best grade exists in such quantities that some of it is out of use all the time, no rent can arise; for the tenant upon whom the demand for

rent might be made would simply substitute some of the unused tract for the one he had been occupying. Under these conditions land is a free good. And since the land is not scarce, the combination which the entrepreneur makes between land and the other factors will be such that the latter will yield the maximum average return per unit of the variables. If, however, land becomes so scarce that the variable factors labor, capital, and entrepreneurship are added in such quantities that diminishing returns set in, then the loss of any amount of land from the combination will force the use of the same amount of labor and capital on a smaller area of land, and the total return from the entrepreneur's operations will be decreased. The entrepreneur then will find that he can afford to pay rent for some additional land that will increase the productivity of the entire combination.

We may suppose that when land is free an entrepreneur makes a combination of 100 acres of land with 50 units of capital and 50 units of labor, and that this combination yields 1000 bushels of grain. This is the best combination. If 90 acres are used, the product will be something less, say 950 bushels. If, now, land becomes scarce, that is, if there are some entrepreneurs without land or some who need more land to make better combinations with their labor and capital, then bidding for the land will begin to take place. The user of the 100-acre tract will either have to give up some land or pay rent for it. If he gives up, say, 10 acres he will find that his labor and capital will produce less grain. If the observed decrease in product is 50 bushels, then the average return for labor and capital will have declined from 10 bushels per unit to 9.5 bushels. Evidently he could as well afford to give the landlord 50 bushels for the use of the land as to reduce his farm to 90 acres, and if there is competition the landlord will be able to obtain rent for the entire tract.

In this illustration nothing has been said as to why land becomes scarce. As far as an entrepreneur is concerned, scarcity arises from the fact that other entrepreneurs are willing to bid for the land. They are willing to do so because they are in possession of labor and capital but without land, or because they have already applied more labor and capital to their holding of land than has the first entrepreneur and therefore can improve the total productivity of their labor, capital, and entrepreneurship by acquiring some additional land.

The same analysis will apply if we assume that a farmer comes from a distant place to begin operations in a well-settled territory. He will find that all the land is owned and in use. We may suppose that he will

go to several farmers and ask them to allow him to cultivate a portion of the land they occupy; but if the land is scarce in the community, if all of it has had applied to it that amount of labor and capital which brings maximum average returns, he will find that none of them will allow him to use any portion of their land unless he pays rent for it. The reason, of course, is that if they parted with any portion of their farms they would be forced to apply their labor and capital more intensively to the areas remaining to them, and, as a result, the average return per unit of these factors would fall.

The scarcity of land arises from natural limitations. And in any given stage of the development of productive technique the land resources available for a given population located in a given geographical area are for practical purposes fixed. Some land can be "made." For agricultural uses the irrigation of arid regions, the draining of swamps, and the diking of rivers and the sea may add small amounts. For urban uses the filling of areas otherwise unsuited for the support of buildings may be resorted to. But these methods can add but a negligible fraction to our natural resources. The more significant method of increasing the supply of agricultural and building land available for the use of the population is through changes in the methods and cost of transportation. Densely populated areas, such as England, Holland, and the northern Atlantic seaboard in the United States, produce in the immediate locality neither all their food supplies nor all the raw materials for their manufactures; they draw on the lands of other and distant regions for these supplies. But their ability to do this depends upon a low cost of carriage. These facts are important when one considers the relation between the amounts of capital and labor on the one hand and land on the other. For many sorts of production the combination between these factors of production must be considered on a world-wide scale.

THE AMOUNT OF RENT

The amount of rent that will be paid for any piece of land depends upon the marginal productivity of that land. Let us assume that a farmer is tilling a tract of 100 acres with the application of 100 units of labor and capital. His crop amounts to 1000 bushels, which sell for \$2000. If he subtracts a small part of the farm, say 5 acres, and leases or sells it to another, he finds that his total product decreases. This follows from the principle of diminishing returns, since the 100 units of labor and capital are now applied to a smaller area of land. If the

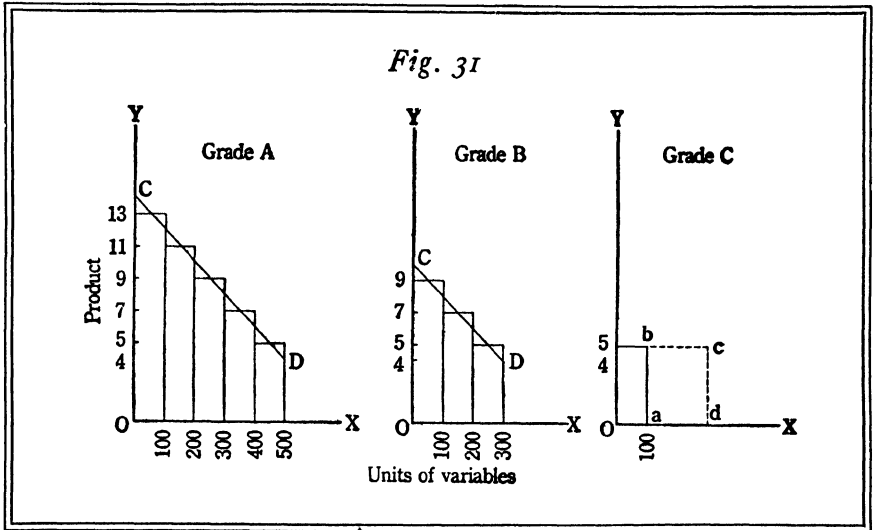
diminution amounts to 25 bushels, then this is the productivity of that small part of the farm. It should be noted that the labor and capital will now produce a crop worth only \$1950. Obviously the farmer would not allow another to occupy this 5-acre tract unless he paid \$50 annually for the privilege. The loss accompanying the subtraction of 5 acres measures roughly the productivity and the rent of that tract. The rent of the entire farm, therefore, is approximately \$10 an acre.

There is, of course, a slight error in our calculations, because 5 acres is a rather large fraction of the entire 100 acres. The subtraction of so large an amount of land might very easily make the use of 100 units of labor and capital unprofitable and cause the farmer to sell or lease some of it to another person. If, however, we should suppose that the subtraction was a very small portion of the farm, then this subtraction would not affect the productivity of the labor and capital (except so far as the land was responsible for that productivity), and the difficulty would be obviated. If 1 acre was subtracted and the total crop reduced by 5 bushels, then the calculation of the rent would be substantially correct.

Farmers probably do not often experiment in exactly the way we have illustrated, but they apply the same principle. They find that rents are fixed by the market and that the prices of labor and capital also are determined in the labor and capital markets. They then make the best combination of the factors, these prices being assumed, and they discover what is the best combination by adding now a little more capital and labor and now a little more land. If, on the one hand, typical farmers find that they can profitably dispense with some land and make a better combination by leasing smaller farms, then the demand for land will decline, and rents will fall; if, on the other hand, they find that with the prices of the factors as they are, better combinations can be made by adding more land to the capital, then the demand for land will rise, and rents also will rise. If all the farmers own the lands they occupy, the same effects will follow; but since no contract rent is paid, the increase in the demand for land will manifest itself in higher prices for it. Each farmer will want more land, and the prospects of bettering the combination of the factors will induce each to bid for additional acres. But the reason for this rise in the price of land is the rise in economic rent.

RENT AND DIFFERENCES IN LAND

In the earliest explanations of rent great stress was laid on the fact that different tracts of land vary greatly in their natural productive-ness. The existence of differences in the productive capacity of dif-ferent areas of land complicates the explanation of rent, but it does not introduce any new principle. The explanation is complicated because the supply of land is made up of competing areas that are not homo-geneous. In Fig. 31 are represented three grades of land, A, B, and C,



which are unequally productive. The variable factors labor, capital, and entrepreneurship are applied to the land in combined units and in doses of 100 units priced at \$10 a unit. Since the prices the entre-preneurs must pay for these factors remain constant, each dose can be measured either in physical units or in money. In the diagrams the physical measure of the variables is employed.

If the demand for a given product, say wheat, is so small that less than 1300 bushels (100×13 , the product of 100 variables on grade A when the best combination is made) is required by consumers, some land of grade A will lie idle. Less than 100 variables will be combined with less than the total area of grade A. Since some of that grade is idle, none of it can have any value, and all land will be a free good. Wheat will sell for its marginal cost in terms of the variables alone, or 77 cents a bushel ($\$1000 \div 1300$); but when more than 1300 bushels, say 2400

bushels, are required, 200 units of the variables will be applied to grade A, marginal cost per bushel will rise to 91 cents ($\$1000 \div 1100$) owing to diminishing returns, grade A will become a scarce agent of production, and all of it will yield rent.

When demand increases still more and reaches a point where it is possible to apply a third dose of the variables to A, it is found that 100 variables will produce as much on B as on A. The total yield derivable from the third application on A, 900 bushels, can be obtained also by cultivating B, and the increase of production by intensive cultivation of the best grade, A, will occur simultaneously with its increase by extension to land of the second grade. The two margins, the intensive on A and the extensive on B, must be equally productive if there is perfect competition and if the variables are mobile, since farmers will not apply more than 300 of the variables on the best land, A, where the productivity of the fourth application of 100 units of the variables falls to 700 bushels, until they have exhausted the possibilities of the second grade of land, which will yield 900 bushels for the first application of 100 variables.

When consumers demand 7100 bushels at \$2 a bushel, the best land will receive five applications of the variables, or a total of 500 units, costing \$5000. It will produce 4500 bushels. On land B 300 units will be applied, resulting in a crop of 2100 bushels. Land C will be the marginal land, because on it \$1000 worth of variables just pays for itself by yielding a crop of 500 bushels, which will sell for \$1000. The price of wheat will now remain at \$2 a bushel until all the land of grade C has been cultivated to the point of maximum average returns.

We now turn to the measurement of rent on the A land. We find that this land now produces, with the aid of 500 units of the variables, a crop of 4500 bushels, which sell for \$9000. How much of the crop will the landlord receive as his share? The answer is that he will be able to compel the tenants to pay over to him the full contribution of his land to the total crop of 4500 bushels obtained from it. This contribution can be ascertained in a number of ways. Suppose that the variables were deprived of the use of land A. Then, in order that the demand might be met, they would have to be employed on the marginal land, where their productivity would be only 2500 bushels. Hence the owner of land A can compel the tenants to pay to him the difference, or 2000 bushels. This is the "bushel rent" of his land. The measurement of the rent of land A can be found also by subtracting the total cost of the variables, which measures their productivity in alternative employ-

ments, from the total value of the common product of these variables and land A. The total value of this product is \$9000, and the total cost of the variables is \$5000. Hence the money rent of land A is \$4000, which agrees with the value of the bushel rent ($2000 \times \$2$, the price of wheat). By similar calculations the rent of B land can be shown to amount to 600 bushels, or \$1200.

The question will probably occur to the student whether the existence of land rent is dependent upon the differences in the productive capacity of land. If we mean by this question, Are land differences as we find them a factor in determining the rent of particular farms and of particular tracts of city land? the answer must be an affirmative one. The existence of poorer land does affect the rent of the better grades, because it offers alternative opportunities for the production of goods. And the rent of any given area does depend upon its superiority to no-rent land. On the other hand, it is not true that rent would not exist if there were no differences in fertility, location, and climate. Let us assume that land A in the illustration is the only land available for the supply of wheat. When more than 1300 bushels are demanded, rent will rise on this land, because the withdrawal of any part of it would reduce the total productivity of the four factors. And when the intensive margin has been extended to the point where 100 units of the variables produce only 500 bushels (demand now requiring a total of 4500 bushels), the rent of the entire area of land A will be 2000 bushels, or \$4000.

In the preceding discussion of the rents of A and B lands, we have assumed that the variables were applied to all grades of land in relatively large separate doses. The areas of the polygons in the figure thus measure the *average* productivity of the successive doses of the variables. In this procedure there is some error. Evidently the first unit of the first dose of 100 units of variables on land A should be more productive than the hundredth unit. The productivity would vary from 14 to 12 bushels. And the marginal productivity of the last unit in the total of 500 units would be something less than 5 bushels. Hence the line *CD* is a more accurate representation of the theoretical operation of the principle of diminishing returns. It is evident also that if the marginal productivity of the variables applied on land A is assumed to be less than 5 bushels, the marginal cost of production will have to be higher than \$2 a bushel, and the marginal demand price will have to exceed \$2. The rent of land A also will be greater than when doses of 100 units of variables were employed. But the principles of explanation

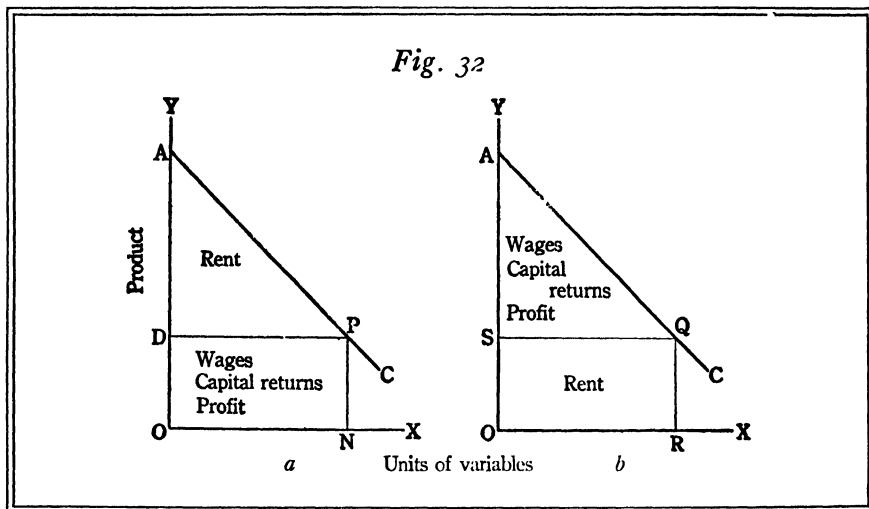
are the same whether the doses are assumed to be very small or relatively large. In dealing with the measurement of rent on individual tracts of land, the supposition of doses of 100 units is probably nearer reality, and it does not lead to an appreciable inaccuracy; but in dealing with the rent of all the land in an entire country, 100 units would be a very minute amount compared with the total capital and labor applied to the land; hence in the explanation of the generalized theory of rent we shall use a continuous curve, such as AC in Fig. 32, to represent the productivity of the variables.

THE GENERALIZED, OR SOCIAL, THEORY OF RENT

Thus far the explanation of rent has been given in terms of individual parcels of land and in terms of the calculations of individual entrepreneurs. To explain the rent that would arise throughout the entire country, let us suppose that the land area is made up of a given number of entirely homogeneous units and that all the entrepreneurs are marginal. Let us regard land as the fixed factor, and labor, capital, and entrepreneurship as the variables. If, now, we apply the existing supply of the other factors—labor, capital, and entrepreneurship—to the fixed area of land, we shall find that as unit after unit of the former is added, the marginal return diminishes. Finally, when all the variables have been applied, a certain product is obtained. In Fig. 32, *a*, this is represented by $APNO$. But the last unit of the variables produces only NP product. The sum of the variables, then, is responsible for the product $DPNO$, and the land is responsible for APD . The latter area in the figure represents rent. But this illustration seems to imply that the rent is a residual share, and this is not the case. In Fig. 32, *b*, are represented the labor, capital, and management as the fixed factor (several factors combined) and land as the variable. When OR units of land have been applied to the other factors, the marginal product is now RQ and the rent is $SQRO$, and the residual area SAQ is the combined product of the factors other than land. The rent of land thus is seen to be determined, as are wages, by the marginal productivity of the factor in question, that is, land.

There is a degree of unreality in the two following illustrations, but it does not arise out of the method of analysis. The lack of reality arises from the assumption that land is homogeneous and that the labor and capital of an entire society are applied to land in "doses" that contain an unvarying combination of these factors.

The illustrations serve one important purpose. They are near enough the truth to show what tends to happen to rents when the other factors increase in amount and the best land in a country has all been brought into use and cultivated past the point of diminishing returns. In Fig. 32, *a*, an addition of the variables past the point *N* on *OX* will result in diminished returns to these variables and an increase in rent. It will have this effect if there is no improvement in technique, so that the



curve *AC*, representing diminishing returns, remains fixed. And such is the tendency. As population increases, the margin of cultivation is forced farther and farther toward the point where no additional raw materials and food can be wrung from the soil and each acre of land becomes more and more important. Its marginal productivity is increased, and the landlord's share in distribution rises.

To combat the tendency to diminishing returns, man has many weapons. But they are all of one type; they all lead to a more nearly complete utilization of resources. We call them "inventions" or "scientific discoveries" or "more efficient labor" or "better methods of agriculture." The effect of these improvements is to raise the curve *AC* in Fig 32, *a*, farther above *OX*. And as long as men can devise new ways of capturing natural forces or of avoiding the waste of those they now hold, population and capital may increase without reducing the marginal product of labor and capital. But it must be remembered that there is an absolute limit on resources, and the law of diminishing returns is *always* in operation.

AGRICULTURAL RENT

The economic rent of any particular farm is determined according to the productivity of the land that is included in it ; but this statement obscures certain details that are of great practical importance. When one says that a certain amount of labor and capital is used in connection with a tract of land, he is ordinarily understood to mean the labor and capital used directly in farming. But production is usually not complete when crops are harvested. Wheat is not a consumable good ; it must be transported to central markets, milled into flour, distributed to consuming centers, and baked into bread before it is ready for consumption.

The labor and capital used in transporting, milling, and baking are just as much concerned in the production of bread as the work done in the wheat fields. That this is the case is known to everyone. But the implications of these facts are not always so well recognized. If much labor and capital must be used to carry the wheat from remote land where it is grown to the consuming centers, the productivity of this remote land is thereby reduced. Such land adds less to the total finished product than that which lies nearer these centers, although it may yield just as large a crop in return for a given amount of agricultural labor and capital. And since its productivity is less, its rent is less. It is customary to speak of differences in rent due to location as *situation* rent. Thus a farm near a large city may be no more fertile than one distant by a thousand miles but may rent for 50 per cent more. This difference is situation rent. All that comes of this distinction between ordinary rent, or *fertility* rent, and the excess return on the more favorably located land is to call attention to the fact that fertility alone does not determine how good or how poor the land in question is. The reason why it does not is that fertility alone does not determine what the marginal product of the land shall be ; something depends on location.

Agricultural land yields various products, some of which are very scarce relative to the demand for them, and others plentiful. This relative scarcity arises from the fact that some crops may be grown on almost any sort of land, whereas others can be produced with reasonable facility only on limited areas. And these limited areas have higher rents than other lands, although they may be capable of producing no more of the plentiful products than the other lands. Lands in the citrus-growing areas of the United States yield rents greatly above the rents of the best wheatland, not because they would produce more wheat, but because they will grow oranges and lemons, and the lands where

wheat is usually grown will not. Still another example is found in the relatively infertile areas near large cities that yield high returns to labor and capital because they are especially suitable, in respect to location, for growing perishable foodstuffs, such as vegetables and small fruits.

URBAN RENT

Land in a city may be used either for residence or for business. The conditions affecting the productivity of land used for these two purposes are not entirely the same. The rent of residential sites is in some respects closely akin to agricultural rent. The demand for a residential site depends on topography, drainage, and other fixed conditions affecting its natural capacity to meet human desires. But man-made conditions affect the demand for urban residential land more than they do the demand for farm land. Proximity to good schools, efficient passenger transport, public improvements, and the fashionableness of the neighborhood are all major factors in the determination of the price that people will pay for the occupancy of city land; they are less important in determining the demand for farms.

The rent of urban land used for business purposes is determined almost entirely by location. Yet the desirability of different locations is largely man-made. Nature does one very important thing; it limits space. Topography plays some part, but not so great a part as in the case of agricultural land. High business rent arises in the centers of the retail and financial districts. Where traffic is great and where many people congregate, it is possible to sell goods more rapidly than in the outskirts of the city. In other words, the limited area in the heart of the retail district has a high productivity because it is possible to transact a very great volume of trade on it. Hence the withdrawal of relatively small tracts would cause a considerable reduction in the total productivity of the labor and capital engaged in production. The withdrawal of the same area of land in the outskirts of the city would not affect total productivity by anything like the same amount. Because it is possible to carry on so much trade on a small centrally located area, lofty buildings are erected and much capital and labor are used in conjunction with the land. It is sometimes said that high rentals are the result of heavy improvements, but this is not the case. One often sees high buildings, erected by an unwise builder, which do not cause the site on which they stand to have a high rent. It is the great demand for the land that causes it to be highly improved and to yield a high rent.

In any city of considerable size rather well-defined zones of use will be found. The highest rents are found in the retail and financial districts, where the concentration of people is greatest and the demand highest. Next in order come the districts of wholesale trade and light manufacturing, then the zone of heavy manufacturing and such businesses as coalyards and grain elevators. Between the business area and the better residential districts there usually exists a borderland that at any given time is devoted to residence. Present rents are likely to be low, but land value may be high because of the prospective use of the property for manufacturing or some other business. Beyond this zone lie the better residential areas, which again are differentiated according to the desirability of the sites.

Since urban rent is nearly all due to location, it follows that the land values and rents in a city are greatly affected by its transportation system. The retail district must develop where large numbers of people can reach it conveniently; the wholesale and manufacturing districts must be served by the railroads; even the residential districts depend in part for their desirability upon the ease with which people who live in them can reach other parts of the city.

THE UNEARNED INCREMENT

Land is the only factor of production whose supply is not affected by real cost. All labor is more or less irksome to the average person; at least, the amount and kind of work performed by a majority of us becomes irksome. Some part of the savings of society is made at a sacrifice of present consumption. Hence wages and interest are thought to be payments necessary to secure the present supply of labor and capital. The real cost of the entrepreneurship stands in a much more doubtful position. It is doubted by some economists that any real cost is entailed by entrepreneurship. On the whole, however, the position of land is unique; there is no vestige of real cost connected with its supply. Moreover, whether real costs are involved or not, it is said that the landlord is in no wise responsible for the supply and cannot affect it. There would be just as much land for use, it is argued, if property in land were abolished tomorrow.

The rent paid to the landlord, it is concluded, is not a payment necessary to bring forth the supply of land. With this conclusion we must in the main agree. Yet there are several objections. The land exists as so much physical stuff. It is not a self-directing agent of pro-

duction. Someone must plan for its use, improve it, and, if it is leased, attend to the business of finding tenants and collecting the rents. Since the landlord performs these functions, the gross rental that he receives is partly rent and partly a composite of wages, capital return, and profits. If there is any excess of returns above the fair payment for the performance of these duties, that payment is true economic rent and is the result of the productivity of the land and not of the landlord's activities. It is unearned and unnecessary so far as maintaining the supply of land is concerned.

So much we may grant. But when it is argued, as it is by the advocates of the single tax, that society should proceed at once to deprive the landlords of their rents, we are not obliged to agree. In the first place, the argument does not take into account the fact that the supplies of some sorts of labor are not dependent on the wages ordinarily paid for them. It is absurd to suppose that the monetary returns received by a great surgeon or by a popular opera star are "necessary" to secure these services. To be sure, the high returns to these professional workers may cause more people to try to enter these fields than if only the wage of common labor were ever earned in them; but that is far from being an identical proposition with the statement that unless the opera star received \$5000 a week the supply of such artistic talent would be reduced. Of more practical significance and therefore of more weight as an argument is the fact that a large number of people receive higher wages than are necessary to cause them to put forth their best efforts in production. Human activity is motivated by a great variety of objectives and conditions. It is fair, therefore, to ask the advocates of the single tax whether they propose to tax away all the unearned increment in wages. To this question at least a partial answer can readily be made. Our interest in high wages is humanitarian and cultural. We believe that for a majority of the people higher wages mean greater welfare. And so long as an augmentation of wages does not, in the long run, check production or cause a decline in the total productivity of all the people, we would not do anything to prevent that augmentation. Rather, we would do everything possible to bring it about. To this answer the opponents of the single tax make the following rejoinder. In the United States the rent of land is very widely distributed. It is not received, as is sometimes said by overardent advocates of its confiscation, by a small, idle landholding aristocracy. Several million farmers receive the bulk of agricultural rents. In the smaller cities residential rents go chiefly to people of the middle classes. The rents of business

property are probably much more concentrated; and in the great cities all rents are likely to go to a smaller, wealthier group. The problem of the "unearned income from land," say the opponents of the single tax, is mostly a problem of the great cities, and not of the country as a whole. Outside these cities rent is part of the income that helps to raise the standard of living of much the same group that is aided by higher wages.

Another objection to the single-tax program, one that runs in the same vein, comes from those who see in the taking of interest by the very rich another instance of "unearned income." That much of the interest now paid is not necessary to secure the present supply of capital might as well be admitted. Not everyone who receives 4 per cent on his investments would cease to save if the rate were lowered to 2 per cent. Since this is the case, some interest, though by no means all, is just as unnecessary as is land rent. However, there is practical difficulty in discovering the receivers of unnecessary interest; there is not the same difficulty in the case of rent. We can determine with a fair degree of accuracy how great the true rent of any piece of land is; we cannot determine whether the interest payment on a certain bond is more than what is necessary to induce its owner to save the money he used to buy it.

It is not the purpose of this chapter to give categorical answers to controversial questions, but to point out some of the fundamental characteristics of rent. However, some conclusions may be stated on the topic of this section. Economic forces cause rent to exist. The institution of private property in land gives all persons the legal right to own land and thus to collect rent. As far as economic reasoning goes, we can safely conclude that rent is an income whose payment to landlords does not affect the total supply of natural resources. But it does not follow that it would be good public policy to confiscate economic rent. Whether or not it would be can be settled only after all the factors affecting social welfare have been considered. Would it be for the best interest of society to take from the present owners of land the rents they expected to receive when they bought land? It may be asked also whether the settlement of America would have taken place as rapidly as it did if the pioneers had known in advance that they would be forbidden to profit from the rise in land values. Would resources be used as efficiently as they now are if rents were confiscated? Would private initiative in the further development of mines and water power be discouraged? These important social problems connected with landownership must be considered carefully before one can arrive at an intelligent judgment as to the desirability of draining rent into the public treasury.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Criticize the following statement: "When the land in the state of X was first taken up it was very fertile. But because of careless cultivation the fertility of the soil was depleted and diminishing returns set in. Now the land is relatively unproductive, products are scarce, and rents have risen."

2. How would an improvement in the transportation of the products of land affect the situation rent (1) of lands near the great consuming centers? (2) of those distant from such centers?

3. How should you expect an improvement that lowered the cost of production on the best grade of land to affect the rent of that land? the rent of the second grade of land, if there are more than two grades?

4. "Rent is a result and not a cause of high price." Is this statement true? Explain why, or why not.

5. "Since man did not make the land, the taking of rent by the landlord is taking something that belongs to all of us. It is a bold, bare robbery, like chattel slavery." Do you agree with this statement?

6. How will the rent of agricultural land in the United States be affected if much of the semiarid region west of the rooth meridian is abandoned? Would this abandonment affect all parts of the country alike? Would California orchards be affected the same as Kansas wheatlands?

CHAPTER XXVI · Interest



Land and labor, the primary factors of production, yield returns that are determined by their marginal productivity. The third factor, capital, also receives a return which, like rent and wages, is a share in the common product of all the factors. This return is commonly called interest; but, as will be explained presently, not all the return to capital goods falls into the category of interest. Part is replacement of the original investment of the owner of the goods. But before we can proceed with the explanation of interest, we must first understand the nature and the function of capital.

NATURE AND KINDS OF CAPITAL

In business usage the term "capital" is given a variety of meanings. Sometimes it is taken to mean the net investment of the proprietors in the business they own, that is, the worth of the enterprise less all obligations to persons outside the proprietorship. In other cases business capital means the total investment of both the proprietors and the bondholders. These two concepts have in common the idea that capital is the amount of the investment in the productive enterprise. But the worth of a business depends primarily on its earning power. If the earning power of an enterprise is so low that no one would give for it the amount of money originally invested, then the earnings of the concern, and not the original investment, determine its value. On the other hand, if the earnings show a return much greater than ordinarily could be obtained from the original investment, it must not hastily be concluded that the concern is worth more than the amount the investors have put into it. We must know whether the surplus earnings are likely to continue in the future; whether they are to be attributed to a specially favorable location, to patents, or to superior management. If we find that the conditions in the business are such that the surplus earnings are likely to continue in the future, we conclude that the capital value of the business exceeds the amount of money the owners have invested in it, and that the business capital exceeds the original cost of the plant.

The economist's concept is different from the businessman's in several respects. In the first place, the economist is not concerned with the total investment. That may consist in part of good will built up by advertising or by custom, which is not capital. Nor is he concerned with the capitalized value of the business as a going concern. For economics the term "capital" means any physical agent of production other than man and natural resources.

All capital is produced by man. It consists of tangible goods, of physical objects that can be produced only by the expenditure of labor and the use of natural resources. Hence all capital costs something to produce. To produce a machine requires a certain amount of labor and a given amount of natural resources. The cost of obtaining these factors is the cost of the machine.

Capital is said to be saved. It comes into existence because of saving. By this we mean that instead of using land and labor to produce goods immediately available for consumption we choose to make machines the product of which will be available only in the future. Saving is the act of setting aside purchasing power for the payment of the expenses of producing capital goods. It involves the decision to forgo immediate consumption and take a money income in the future instead.

Producers' capital consists of all produced means of production employed by entrepreneurs in business undertakings; consumers' capital consists of produced means of production that yield utilities directly to their users. The distinction is not made on the basis of any physical difference between the two kinds of capital but on the use to which the capital is put. A house owned by the person who lives in it is an item of consumers' capital, and the same house used for a shop or a workplace is producers' capital. For the most part, however, there is a physical difference. The greater part of producers' capital consists of factory buildings, machines, railroads, and commercial structures. There is, however, a more important distinction than difference in use. Consumers' capital is always used for the direct gratification of personal wants. Producers' capital, by contrast, is typically employed by entrepreneurs to produce consumers' goods or other capital in order to obtain a money income. Its employment is controlled by purely business motives, and it usually is valued for the monetary income it will bring in. The employment of consumers' capital is affected by whim, fad, impulse, and all the other imponderable considerations that rule in the domain of consumption. In the discussion of interest we shall confine our attention to producers' capital because it is in respect to

such capital that the phenomenon of interest becomes most important. There is no controversy, for example, about the earnings of consumers' capital. Labor does not insist that consumers' capital receives too large a share of the national income, but it often does make that charge concerning the earnings of producers' capital.

It is not usual for an economist to call a nondurable commodity consumers' capital. Bread is not capital but a *product* or a *commodity*. Sometimes it is said that this distinction rests on the fact that bread is not durable. But the coal used in a factory ordinarily would be called producers' capital, although it is consumed almost as quickly after production as bread and many other types of consumers' goods. Evidently there is a lack of consistency in the application of the term "capital." Generally speaking, consumers' capital embraces only durable objects of consumption such as houses, furniture, carpets, and automobiles used for pleasure. But producers' capital consists of all material wealth that aids in further production.

KINDS OF PRODUCERS' CAPITAL

Producers' capital is not homogeneous. Some forms are durable and others are perishable; part is specialized and part is nonspecialized, or free, capital. A machine designed to perform a mechanical operation of a limited sort is called specialized capital; a building that can be used for a variety of purposes is called free capital. In general, most raw materials, metallic money, some machines, and some types of buildings are examples of free capital. The entrepreneur can *withdraw* his free capital quickly from an unprofitable employment, but not his specialized capital. So, too, he can withdraw his investment from perishable capital, but not from durable capital. From this difference arises the fact that old specialized capital may yield a net return that is far below the going rate of interest. It continues to do so because it cannot be shifted to any other use, and because it has not entirely ceased to be productive and hence the entrepreneur cannot afford to abandon it.

In Chapter XXIV we used the term "mobility of labor." Capital also possesses a certain degree of mobility. Free capital has industrial mobility; that is, it can be diverted from the industry toward which it was intended to move and be directed to other uses. But much capital has not a great deal of mobility; it must remain where it is first installed, and most of it must stay in the industry for which it was made. Yet we constantly read that capital has been shifted from one industry to an-

other, as though it could be moved about with the greatest freedom. What usually happens is that investments have been shifted. Capital yields its owners a return that they may invest in a great variety of undertakings. A part of this return, as we have already said, is repayment of the investment. The machine wears out; but if the investment of money in it has been wise the owner will receive from it during its lifetime a sufficient amount to pay back its original cost plus interest. He may use the interest for the purchase of consumers' goods; but unless he wishes to dissipate the wealth he owns he must reinvest that portion of the earnings of the machine which amounts to its original cost. This we may call the replacement fund.

When capital shifts from one industry to another, it is the replacement fund that is diverted. As shipping declined, capital was shifted to the textile industry. No diversion of capital goods could have occurred, of course. What probably happened was that the owners of vessels failed to replace them and diverted the replacement fund to the purchase of factory equipment. In the same way, investments in perishable capital are constantly on the move. In one year the banks may lend large sums of money to one industry and in the next to an entirely different industry. Because bank loans are commonly used to purchase raw materials and to pay laborers, they are repaid quickly. When the demand for loans in one line of production declines, the banks seek openings in another quarter; hence the capital goods that the banks finance may be greatly changed in a very short period of time.

It should now be clear that capital is not a "sum of values." It always consists of physical objects. But these objects have value; and since they are very unlike physically, about the only way in which they can be reduced to a common denominator is in terms of value. It is for this reason that the businessman, when asked how great his capital is, will usually reply by stating its estimated money value. But the money value of capital is no more capital than bushels are wheat.

THE RETURN TO CAPITAL GOODS

Capital, as a physical agent of production, yields a return based on its marginal productivity. That is, a machine is a factor of production as much as the labor that tends it or the land on which it stands. Like that of the other factors, its return is governed by marginal productivity. The marginal entrepreneur discovers the productivity of capital, just as he discovers the productivity of land or labor, by varying the propor-

tions of all three factors until he obtains the combination that will produce goods at the lowest unit cost. If marginal entrepreneurs find that they can profitably substitute labor for capital, then there will probably be a surplus of the latter; its marginal productivity and price will decline, and old capital will be revalued. With this fall in price more capital will be employed, a resubstitution will occur, and finally the total supply will be taken into industry at prices that correspond to the productivity of capital.

Unless the price paid for capital covers its cost, its supply will be contracted and its marginal productivity will rise. Hence there is a relation between the cost of capital and its earnings. But this relationship is very complex. In the first place, it should be noted that the entrepreneur makes a different calculation for the machine from the calculation he makes for labor. The product of labor is work, and for that the entrepreneur pays wages. He does not buy the laborer; he does buy the machine. Having invested a sum of money in it, he must look forward to replacing it when wear and tear and the progress of technique have rendered it unproductive. He does not need to make any such provision for labor. Moreover, he buys the services of labor and, within a short time, turns these services into goods, which he sells. But when he buys the machine, ordinarily he must wait for his return and must take risks that come with decisions in respect to future events.

In buying a machine the entrepreneur must first estimate its marginal productivity, which he then translates into its annual earnings throughout its life of usefulness. From these earnings he then must deduct an annual amount which, when accumulated during the life of the machine, will return to him the original investment. This accumulation is called the replacement fund or, in accounting terminology, the reserve for depreciation. If the machine does not earn this amount, it is clearly an unprofitable investment. And it must earn something more than this if the entrepreneur is to receive any interest on his savings invested in it. For example, a machine that cost \$1000 and which wore out in one year would need to earn at least \$1000 during the year before the entrepreneur owning it could recover his original outlay. If it earned \$1060, and if the going rate of interest was 6 per cent, then its marginal net product of \$60 would exactly equal the interest charge for that time. To put it another way, a machine that would earn \$1060 gross would be worth \$1000.

If men were willing to exchange *present* goods (goods whose money value or consumers service can be had *now*) for an equal amount of

future goods of the same sort, then the producer of a capital good could sell it for the sum of its marginal gross products. And such capital goods would be multiplied until their costs and the marginal products were equal. But, as a rule, no exchange between present and future goods takes place unless the future goods have first been scaled down in value or discounted at the going rate of interest.

We may summarize the discussion of the productivity of capital thus far given as follows: Capital goods, like other factors of production, earn returns based on their marginal productivity. But, unlike labor and land, the supplies of capital goods are affected by their direct money costs, and unless these costs can be defrayed from the marginal product of the capital goods they will not be produced. Partly because they are durable, but chiefly because the capitalistic process of production requires time, capital goods yield their products mainly in the future. Since men do not exchange present for future goods unless the latter are discounted, capital goods ordinarily must yield returns that exceed their costs. If they do not, the labor that might have been invested in them will be diverted to the production of goods and services for more immediate use.

THE SUPERIOR PRODUCTIVITY OF CAPITAL

It is evident that capital labors under the handicap of having its returns discounted at the going rate of interest. The question arises, then, why men ever employ capital when they might easily use labor and land for immediate production. The answer to this question is found in the superior productivity of the roundabout, or capitalistic, method of production. A thousand days' labor applied to the making of machinery will often yield many times the product that the same quantity could create if it were applied directly to natural resources. For example, a savage can grow scanty crops without the aid of a plow or any other capital. He merely scratches the soil with the first conveniently formed stone or stick that comes to hand, and scatters seed. A somewhat more highly developed civilization makes use of crude implements and gains a much larger harvest. Modern agriculture utilizes tools that can be had only as a result of a high degree of development of the arts of metallurgy and engineering. Instead of setting to work to produce the crop in the shortest period of time and by the most direct methods, capitalistic production engages in many preliminary processes before any consumable product emerges. As the technique of produc-

tion becomes more and more complex, these steps are multiplied in number, and the time that must elapse between the inauguration of the process and its completion becomes longer. The longer period of time is made necessary by the variety and number of the preliminary steps and by the fact that the machines and other equipment are durable. They are durable not by choice, but because there is no escape from durability. Here we must guard against a possible misapprehension. Men contrive to the limit of their powers to shorten the time required by the productive process. Factory-owners no more desire to have their capital tied up in raw materials and half-finished goods than merchants want to keep goods on their shelves a long time before they sell them. They do not prefer a machine that turns out ten thousand articles in ten years to one that shapes the same number of articles in five years. But unfortunately such capital goods as machines cannot quickly be exhausted of their productive capacity, and as a consequence a capital-using industry must usually put more time between the beginning and the end of the productive process than the direct, noncapitalistic industry does.

Since the machine method, or capitalistic method, is always a roundabout, rather than a direct, application of labor, it is sometimes said that roundaboutness increases product. This is not the case. Roundaboutness is an unescapable incident of the method, but it adds nothing. The reasons for the superiority of the roundabout, machine, or capitalistic, method are not difficult to discover. The machine method is more productive, in the first place, because it enables man to capture and control natural forces that otherwise would escape him. To the savage the waterfall is not an aid in production, because he cannot utilize its energy; nor could modern man were it not for the use of capital. In the second place, the capitalistic method enables man to accomplish certain results that are impossible for unaided labor: to tunnel mountains, to create waterfalls by means of dams, to move great weights, and to move weights at a speed beyond the capacity of direct labor or of work animals. There is no other reason than these two. The fact that the method is roundabout does not increase the productivity. Some productive processes seem to require merely time for their completion. When wine ages it changes in quality and therefore in value. Fruit ripens and lumber cures. Behind these common terms, "age," "ripen," and "cure," lie chemical and physical processes. And these happen in time; but it is not the passage of time or the roundaboutness of the processes that produces the change.

Not all roundabout processes are superior in productivity. Many ways of spending labor in the making of machinery cannot be employed because they are either not superior to labor working more directly with natural resources or they are not sufficiently superior to offset the discount rising from the existence of interest. The entrepreneur selects from among all the various methods of using labor, both direct and indirect, those that yield the maximum returns. The rate of interest prevailing in the community and the productivity of the roundabout method, as compared with the direct method, both affect his decision.

TIME PREFERENCE

We have said that men generally will not exchange a present good for a future good unless the latter is more valuable. This general statement is only typically true. We must note first that the statement does not say that an individual prefers *any* present good taken at random to a given future good. It would be absurd to suppose that he preferred a present good of low utility to a future good of high utility. The statement does say that he usually prefers a present good to a future good of the same sort. Moreover, the statement does not take account of certain cases where desires change with the passage of time. The school-boy has no present use for the paraphernalia of a lawyer's office. He knows, if he intends to become a lawyer, that he will require an office in the future, and he prefers to have it then and not now. In winter one does not usually want ice as much as he will want it during the coming summer. These and similar cases are modifications of the general statement. It is admitted that individual valuations change with the passage of time, and the individual knows this fact and takes it into account; it is admitted also that individual valuations change with changing environment, and this too the individual knows and takes into account.

The explanation of this customary discount on future goods is not entirely clear. Several reasons for it have been advanced by different economists, and we may note them briefly, with the warning that all are largely conjectural. In the first place, the fact that men know the future to be uncertain is said to be a factor causing the discount. Since they may not live to enjoy future goods, they set a lower value on them than on present goods. There is also the chance that conditions will change and that the future goods may not yield as much satisfaction as they now anticipate. It is also said that men, being generally optimistic, believe that they will be better provided with goods in the future than they now

are and therefore, according to the principle of diminishing utility, scale down the value of future income.

Whatever may be the cause of the discount on future goods, we are sure that it exists, because the evidence lies everywhere about us. Men do not lend money, except as an act of courtesy between friends or out of charity, unless the lender agrees to return the original loan with something in addition. A child offered a choice between a treat today and a treat a month hence usually chooses the present enjoyment. This phenomenon of discount of the future has been called "time preference" by Professor Irving Fisher.¹ In his words, "It is the (percentage) excess of the present desirability of present goods over the present desirability of an equal amount of future goods." In a generalized statement time preference is a preference for *present income* as compared with *future income*; and in terms of the money economy it is a preference for *present purchasing power* as compared with *future purchasing power*.

Individuals differ considerably in respect to their preference for present rather than future income. Some can scarcely save anything at all, so great is their preference for the present. Others constantly deny themselves present enjoyments for the sake of piling up investments from which they hope to derive a future income; and they do this to such an extent that, in the opinion of their friends, they live constrained and unpleasant lives. Why these individual differences in time preference exist does not concern us very much. Like the variety of tastes and desires for present consumers' goods among different people, they are things that a preliminary study of economics takes for granted.

The same person may have different time preferences under different circumstances. In youth most men are prodigal of income; in middle life they become more parsimonious. But the principal external cause of individual variations in time preference is difference in wants and provisions for wants, that is, difference in the amount of individuals' income as compared with their needs. A poor man can hardly be induced to save anything at all. If he saves he must forgo some of the necessities of life, and scarcely any rate of interest could compensate for that. The rich man, on the other hand, may easily save a large portion of his income.

The rate of individual time preference is affected by the future relationship of income and needs as well as by the present relation between

¹ Irving Fisher, *The Rate of Interest*, p. 88. Copyright (1907), by The Macmillan Company. Reprinted by permission.

these two sides of the personal budget. Man anticipates and plans for the future, and he therefore takes into account in his activities and decisions many future events. For example, if income is rising and needs in the future are expected to remain constant, the relation between income and needs in the present will be affected. Under these conditions the rate of time preference tends to be high, and one may reasonably borrow to increase present consumption. But if future needs are expected to rise and income to remain constant, the individual tends to have a low rate of time preference and to save something to eke out his future income. Thus in the private economy of those who have more than enough income to buy the urgent necessities of life, there is a constant attempt to balance present expenditure and income with future expenditure and income.

THE MOTIVES FOR SAVING

Saving takes place because men wish to realize future objectives. In some instances these objectives are very definite. The saver wishes to spend a year in travel or to buy an expensive consumers' good that he cannot buy out of current income. Savings banks know that this is a common motive for saving and make an appeal for deposits for the purpose of buying Christmas gifts or for financing vacations. In many other instances the objective is not clearly defined. Many people save for such general objectives as security in old age, the proverbial rainy day, and the enlargement of their income and their scale of expenditure in the future; still others save to build an estate for their children or to enlarge the size of their business operations. To say that time preference exists is not, of course, to deny that men desire future income as well as present income.

Some saving undoubtedly would take place if no interest were ever to be realized from it. Needs are not necessarily distributed throughout life as income is. If interest disappeared entirely people would still wish to provide for future needs. The parent saves while his children are young to be able to expend more for their care when, a few years later, their needs have become greater. In many instances he would save even though no interest could be got from money deposited in the bank or invested in bonds. To do this he might hoard consumers' goods, but not to a very great extent, since commodities are usually subject to depreciation with the passage of time. Metallic money, plate, and jewels are not subject to physical depreciation to the same extent as

other commodities, and so we find that in backward countries, where opportunities for investment are few, saving generally takes place through the hoarding of these goods.

It is sometimes said that the rich would save even if no interest could be got from investment, because the act of saving involves little or no sacrifice on their part. There are objections to this assertion. In the first place, the accumulations of the rich are largely derived from the net income of capital or from interest itself. If interest were cut off, the savable surplus of the owners of great wealth would be diminished and saving would decline. In the second place, the statement overlooks the fact that the rich could spend a great deal more on immediate goods than they now do, though they might not spend a very much larger amount on personal gratification of the meaner sort. In Oriental countries large amounts are expended by the rulers and the rich on the maintenance of personal retainers, on the building of luxurious palaces, and on other forms of display. In the Western world the same sort of consumption takes place, but to a relatively more limited extent. This contrast is doubtless due to obscure differences in the ideals and customs of the people of the two regions, but no small part of it is due also to the opportunities for the earning of interest on capital which accompany the capitalistic regime in progressive countries.

Under the social conditions prevailing in the United States, it may be true that the very rich would continue to set aside for future production the larger part of their receipts from rents and profits even though interest were considerably reduced. They would do so partly because of habit and partly because business has become for many of them an end in itself, and saving largely involuntary. We must be careful, however, not to conclude that they would continue to save as much as they now do. The habit of saving and the motives that now influence the accumulation of capital are not inborn characteristics of men and might easily be destroyed by an attempt to abolish interest.

Another form of so-called involuntary saving, which is perhaps more important than the personal accumulations of the rich, is found in the building up of large surpluses by corporations. The net earnings of a corporation are legally at the disposal of the board of directors, who may either disburse them as dividends or reinvest them. At the present time a large fraction of all national savings results from such accumulations. Here, we are told, is a case in which the decision to save is often made by a group, the directors of the corporation, who do not own more than a small proportion of the amounts thus accumu-

lated. But the corporate surplus often belongs to many small stockholders who have no voice in the determination of the policies of the corporation. It is said that these stockholders might not save the net earnings of the corporation if these earnings were disbursed as dividends, and that the retention of them as corporate surplus compels the small stockholders to save whether they wish it or not.

When knowledge of the fact that a surplus is being plowed back into the business is kept from all but the inner circle of the management, the argument and the conclusion are of some weight. If the stockholders do not know that the saving is taking place they cannot, of course, take any action concerning it. But there are objections to the conclusion that a lowering of the interest rate would not affect these accumulations. In the first place, if the rate of interest were much reduced, the savable surplus of the corporations would be reduced; in the second place, the directors cannot deceive the stockholders forever, and at some time the facts must be made known. Then the policy of the corporation will be changed if the stockholders do not approve it.

If we assume that the accumulation of the surplus is known to the public, then we must suppose that a majority of the stockholders are in agreement with the saving policy, and the fundamental difference between corporate savings and individual savings is not great. Probably the corporate method results in somewhat larger accumulations, since having money on hand is a different thing from having an undivided share in a corporate surplus. It is easier for the individual to spend the former. The small stockholder feels that he cannot affect the policy of the corporation, and he makes no effort to do so; but if he is greatly dissatisfied with the decision of the directors to retain the surplus, he may sell his stock. If the new owner knows the policy of the corporation, his purchase is evidence of his willingness to have the profits retained in the business.

No one saves all his income. Necessary living expenses must be met first. All in excess of this is savable surplus. Given a motive for saving, the average person then sets aside a certain sum, the size of which depends upon (1) the going rate of interest, (2) the urgency of his desire for the object for which he saves, and (3) the rate at which the marginal utility of the income remaining for present consumption rises as more and more is diverted to future uses. At some point another \$100 added to savings makes possible a future objective that just balances in desirability with the desirability of the present goods the \$100 would buy. At this point the saver ceases to add to

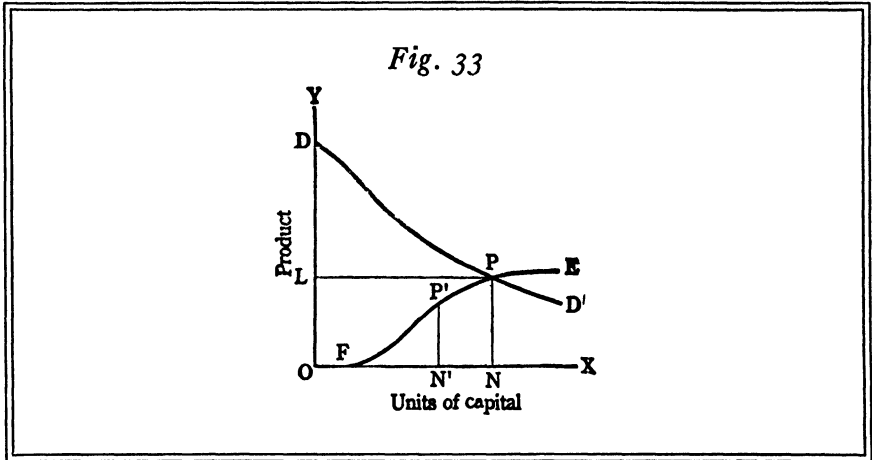
savings and consumes the remainder of his income, and at this point his rate of time preference is equal to the rate of interest at which he can invest his savings.

THE SUPPLY OF SAVINGS

The total supply of savings, that is, the supply of current income offered for financing the building of new capital, is made up of the supplies of numerous individuals. As we have seen, the amount any individual will offer varies with the rate of interest; and the amount the entire body of savers will bring to the loan market also varies with the interest they can get, just as the total supply of a commodity varies with the price that the consumers are willing to pay. Like the supply of any commodity, the supply of savings may be affected by conditions other than the interest rate (price in the case of a commodity). Changes in the attitude of parents toward their children, changes in the standards and modes of consumption, changes in the distribution of the national income, all have some effect on the price at which different parts of the total annual savings will be brought to the loan market. At present we are concerned with the relation between the rate of interest and the volume of savings.

Some part of the supply of savings would be forthcoming even if no interest were paid, and some if a substantially lower rate than the one prevailing were offered. But the going rate of interest is necessary to bring the supply of savings up to its present volume. The supply of savings may be illustrated by the curve FE in Fig. 33. In this figure the curve DD' represents the diminishing net productivity of capital; that is, it represents the demand for the use of savings, which arises out of the existence of the land and labor used in production, just as the curve AC in Fig. 29 on page 385 represents the demand for labor. The demand for savings will be discussed in the next section. Here we are concerned with the cost and supply of savings. If only a very small amount of savings were demanded, it could be obtained at zero interest. If ON' units are to be supplied, $N'P'$ interest is necessary; but if ON units are required, NP interest is necessary. And the rate that would correspond in percentage to NP is the marginal cost of savings. The area $LPNO$ is the extra amount of future goods that must be paid in an exchange of present goods for future goods. For since all savers are paid at the same rate, not only the marginal increments but also all other portions of the annual total of savings will receive NP .

It must not be supposed that the curve FE represents the interest rates at which different persons or groups of persons can be induced to save; it shows the cost of calling forth different amounts of savings from the entire community. One person might save a small amount at zero interest, a larger amount at $N'P'$, and a still greater amount at NP ; for others the minimum effective rate might be $N'P'$ and for still others NP .



The marginal cost of saving is not a fixed, or constant, amount. Changes in the income of the mass of savers will affect it. If spending habits remain constant and if income increases, the rate which would call forth a given volume of savings will fall. If, on the contrary, income diminishes, the rate will rise. Wars and other agencies that destroy capital or reduce income tend to raise the marginal supply price of savings. Changes in the habits of consumption of a people also will affect the time preference of a majority, a tendency to easier living and spending being accompanied by rising rates and a reverse tendency by falling rates. Thrift can be taught, and there is no doubt that an intelligently conducted campaign for the saving habit would bring results. Unthrift also can be taught, of course, and every seller of luxuries who, by advertising and other means, persuades people to buy his wares is engaged in teaching it. The government may also contribute to the decline of saving by taxing away the savable surplus of the well-to-do and spending the proceeds of the taxes on consumable goods. It may do the same thing by weakening the motives for saving and by increasing the uncertainties attendant on investment.

THE DEMAND FOR SAVINGS

The demand for savings arises from two sources. In the first place, many persons want loans to buy consumers' goods. They want an automobile, but they have not the funds in hand to pay for it. Sometimes such installment purchasers can reasonably anticipate a larger income in the future, but more often they wish to spend their income before they have it. For this privilege they are sometimes induced to pay very high rates of interest. As was explained when time preference was discussed, they discount the future at a high rate. It is obvious that consumption loans present a different problem from loans used in production. Like the employment of labor for direct personal services, the consumption loan involves no allocation of a joint product among the several factors of production. Just as the prices one pays for the services of a caddy or a hairdresser depend on the importance he attaches to money on the one hand and the significance to him of the services on the other hand, so too the interest one pays for the privilege of having a new car a year before he has the funds to pay for it depends on the personal evaluation of two things, the importance of the additional money (carrying charge or interest) he will have to pay and the desirability of having the car a year sooner. In both cases the principles of value rather than those of distribution are involved.

In the second place, entrepreneurs want present purchasing power to invest in capital goods. This is by far the most important demand for loans, and it is the one that is predominant in determining the normal rate of interest. The demand of the entrepreneur for loans is based on the productivity of capital. To return to an earlier example, if the entrepreneur can buy for \$1000 a machine that in the course of its life will earn enough to repay its original cost and at the same time yield a surplus of \$60 a year, then he can afford to offer 6 per cent for a loan of \$1000 in order to make the purchase. That is his maximum demand price for a loan that enables him to buy the machine.

If we should take a limited view of the loan market, we might conclude that the explanation given in the preceding paragraph was a satisfactory account of the demand for savings. But this explanation assumes that the machine which will earn replacement and \$60 a year can be bought for \$1000. Now the value of any producers' good depends on (1) the value of the products, (2) the rate at which the value of the products expressed in money is discounted, and (3) the distribution of the products in time. The value of the products of any capital good

depends, as we have seen, on the demand for them and on the amount brought to market. But the product of the machine is not a separate product, and the value that the entrepreneur will impute to the machine depends on the marginal productivity of the machine. If any kind of capital good becomes more plentiful, its marginal productivity will decline, and the entrepreneur will pay less for it.

The money cost of any item of capital goods (say a machine) arises from the fact that there are rival demands for the factors used in making it. Because of these rival demands the owners of the factors can exact a payment for their use in making capital goods. Even if interest did not exist, and if capital goods earned only enough to replace themselves, the law of opportunity costs would operate. But when the investment in the various sorts of capital goods must earn interest, it is evident that the earnings of these goods must exceed their costs. In Chapters XXIII and XXV it was explained that the marginal productivity of labor and of land determined their earnings. The same principle governs the earnings of concrete capital goods.

But at this point another fact must be taken into account. Although capital goods are the products of the primary factors of production, their productivity is not the same as the productivity of these factors. The capitalistic method, as we have seen, increases the productivity of the primary factors, and the business demand for savings arises from this fact. If savers can be found who are willing to wait for the products of the machine and allow labor to enter into the roundabout method, more products or more valuable products can be created than if all the labor is used directly. And as long as there are methods of attack on the environment that give better results in terms of value than direct methods do, entrepreneurs will demand capital goods in order that they may take advantage of these methods; and as long as *new* methods can be discovered, the demand for additional capital will continue.

As machines and other capital goods are multiplied, the marginal productivity of each machine becomes less and less, following the universal law of diminishing returns. To put it in another way, as the labor and land that go into the machine are increased in amount, they become less and less important as compared with the labor and land used directly. Not all capitalistic processes are equally productive. Some excel the direct method to an almost unbelievable extent; others add but little. In any given stage of the development of technique the most productive are selected first for exploitation. But finally a point is reached where the multiplication of capital goods of the more pro-

ductive sort has gone so far that the amount by which the machine method excels the direct method is just sufficient to offset the interest charge which the savers require to induce them to wait for their income. At the same time that these machines are being multiplied, other less productive machines and capitalistic processes will be brought into use in industry. Many of these will not be as productive as others. There will be all degrees of gain over the direct method. We may call these latter processes inferior, just as we call the less productive areas of land inferior. The decline in the additional product from these increasingly inferior machines marches along with the decline in the productivity of the additional machines of the superior sorts. Both lines of development finally reach the point where the superiority of the capitalistic process is just sufficient to offset the required payment for waiting.

In terms of Fig. 33 the demand of all society for savings is illustrated by the curve DD' . As savings increase, the purchasing power that is turned over to the entrepreneur is invested in capital goods of less and less productivity, indicated by the slope of DD' , until the marginal unit at NP just yields a net return equal to the interest that must be paid to call it forth. The curve DD' in this figure does not represent the gross productivity of capital goods, but only the excess productivity of the capitalistic method over the direct method. The marginal unit of capital goods must yield, in addition to its cost of replacement, a surplus measured by NP .

At this point new savings tend to decline, for further accumulation would not pay the premium demanded by savers in the exchange of present goods for future goods. Whatever quantity of capital goods has been brought into existence at this point is the amount that labor and entrepreneurship will have to work with. It is also the amount that will be applied on the land. The marginal cost of savings limits the supply of capital goods. And that supply in turn determines the marginal productivity of capital goods. It partly determines the wages of labor, the earnings of entrepreneurship, and the rent of land, but only partly, because the productivity of these factors too is affected by the supply of them. The earnings of labor, land, capital goods, and entrepreneurship are mutually determined; that is, wages depend on the marginal productivity of labor, but how much a given number of laborers can produce depends on how much land, capital goods, and entrepreneurship are combined with them. The rent of land is also dependent on the supplies of labor, capital, and entrepreneurship that are combined with the land. Given the state of technique in produc-

tion, the productivity of each factor depends on its supply and on the supplies of the other factors.

We now return to the question, What determines the price of the machine purchased by the entrepreneur? If we are concerned with a single machine or type of capital, the answer is that the machine must sell for its cost of production, including the wages of labor, the rent for land, the earnings of entrepreneurship, and the returns to other capital; but if we substitute for a single type of capital *all capital*, the answer is that the cost of capital goods is simply the value of labor and entrepreneurship working on land without the assistance of capital. This is not entrepreneurs' cost; for entrepreneurs do not deal with capital in the mass, but only with particular units of capital goods. But from the social point of view the cost of production of capital is as has been stated. This explanation at first seems incorrect, because we know that without the assistance of capital goods labor could produce very little, and the cost of capital therefore would seem to be much too small. But as labor is withdrawn from direct production the marginal productivity of the remaining labor rises, and the productivity of the labor embodied in capital goods falls as more and more of it is put into them and the capital goods in which it is invested are multiplied. A balance is reached when a further subtraction of labor from the more direct methods would destroy a marginal product that was greater than the marginal product added by capital discounted to the present.

THE CAPITALIZATION OF INCOMES

When the income from any durable agent of production is reduced to present worth, that income is said to be capitalized. A tract of land, for example, the annual rent of which is \$100, is worth \$2000 when the interest rate is 5 per cent. This is the capitalized value of the annual rent forever. It would be worth more than that if the rate of interest were lower, and less if the rate were higher. The income from the capital goods also may be capitalized; but since these goods are seldom as durable as land, and since they must all be produced, the phenomenon of capitalization does not play as important a role in determining their values as it does in the case of land. If the capitalized value of the future earnings of machines seems unlikely to amount to the price asked for them, the entrepreneur will not purchase. If the capitalized value of the machines is in excess of the price, the demand will cause the price to rise; but the rise in price will be answered by an increase in supply;

the added supply of machines will reduce their marginal productivity, earnings will decline, and the capitalized value of the machines will approach their cost of production. The rate of interest thus acts as a selective factor in determining which machines or other capital goods shall be produced.

When capital has once been installed, it will continue in use as long as it has any marginal productivity. This statement needs some explanation, because we know that often capital goods are discarded before they have ceased to be useful, and better equipment is substituted. In these instances the capital is not really productive. It is either so far worn out or so antiquated that it detracts from the productivity of the labor and land with which it is used, and cannot be said to earn anything. Its value is solely the scrap value of the material in it.

Although capital goods may continue in use as long as they have any marginal productivity, old capital goods may yield a return that is much below the amount required to replace the original investment in them and pay the going rate of interest. These items of capital yield a return that, like the rent of land, is not related to their cost of production. To illustrate this statement we may suppose that an old machine yields \$200 a year, which exceeds the annual replacement charge by only half the amount necessary to pay the going rate of interest on the undepreciated portion of the original investment. The machine's value is now determined by its earnings, whatever they may be for the period of time it may be expected to remain in use, capitalized at the going rate of interest. The original cost of the machine less the reserve that its owner has set aside for depreciation may greatly exceed the capitalized value of its earnings. In the same way, the value of capital goods that have been installed by incompetent entrepreneurs does not depend on the cost of those goods. The original investment in those goods no longer yields the going rate of interest. Strictly speaking, only new, well-planned investments are likely to yield the current rate of interest in a dynamic society. Old and badly planned investments have a return that resembles the rent of land. It is conceivable, of course, that old capital goods might yield a return in excess of replacement or depreciation and the going rate of interest. But this condition cannot persist very long. If a machine costing \$1000 yields something more than replacement and the going interest charge, the supply of such machines will be increased, the marginal productivity of each will decline, and its earnings will fall to the amount which when capitalized is equal only to its cost.

TYPES OF INTEREST

We have used the term "interest" to mean the payment necessary to equalize the values of present and future goods. The typical example is found in the case of a loan to a well-established government. Here the lender turns over to the borrower a certain sum of money and is assured by his contract that at stated intervals he will receive interest, and that at the expiration of the term of the loan his original investment will be returned to him. There is a minimum of risk attaching to the loan, and the conditions of the contract are clear and explicit.

The investment of the entrepreneur in capital goods does not involve a contract between a borrower and a lender, and yet the investor expects and ordinarily receives payment for waiting. This is called *implicit* interest to distinguish it from the payments made on a loan contract, which are called *explicit* interest. The investor in stocks may receive two types of return, implicit interest and pure profits. The former is determined according to the demand for waiting and the supply of waiting; the latter, according to the principles governing pure profits. In numerous cases the common stockholder receives nothing at all. But the same may be true of those who buy bonds or lend on other security. All investments involve a risk that neither interest nor principal will be repaid. To cover this risk the investor requires an addition to the riskless, or pure, rate of interest. When government bonds selling at par yield only 3 per cent, the bonds of corporations whose future earnings are insecure often yield 5 or 6 per cent. The addition of 2 or 3 per cent to the riskless rate is to cover the losses that the investor believes will occur, on the average, in respect to the risky securities. If these estimates or risks were accurate, then a large sum of money invested in government bonds and the same amount invested in poorer bonds would, in the long run, bring in about the same amount of interest. But the average investor cannot distribute his holdings of the poorer bonds widely enough to obtain the benefit of the law of averages. Hence if he is lucky he may receive more interest than if he had bought government bonds; if he is not lucky he may receive less.

Some investors cannot afford to take great risks. Their savings are small, and the consequences of loss are too serious. Other investors can afford to take greater chances. The supplies of savings are not homogeneous. It follows that the interest rates yielded by very secure bonds and by riskier securities may differ by more than the true insurance premium.

THE LOAN MARKET

In the preceding sections of this chapter we have refrained intentionally from introducing any discussion of the manner in which savings find their way into the hands of entrepreneurs, and we have not dealt in terms of market rates of interest. We have been concerned almost altogether with the problem of normal tendencies. When an individual saves, he refrains from spending all his income on consumers' goods and invests the remainder in some form of capital. He may do this by actually buying the capital for use in his own business. From the gross receipts of production the manufacturer must pay all expenses of production. If a surplus remains, he may either spend it all for consumers' goods or buy more equipment or more land. If he buys land he merely displaces some other owner; if he buys more equipment he causes more of these capital goods to be produced.

In many instances the saver cannot invest directly by buying capital goods. He must rely on the loan market for his investments. Having a savable surplus, he may take life insurance, buy bonds, or merely leave the money with a bank. In any case purchasing power is now available for investment in new capital. But the saver does not make any calculation as to the probable future return from capital. The entrepreneur takes the responsibility for choosing the capital goods that shall be bought with the purchasing power that the saver does not wish to spend on consumption. He makes a contract with the saver to pay him a stipulated rate of interest and to return the principal of the loan at a definite date in the future. Sometimes the investor buys stocks and not bonds. But the investor in stocks does not determine the kind of physical good that shall be added to the equipment of the corporation. The managers of the corporation, acting as his representatives, do that.

When one saves, he causes purchasing power to be used to produce capital goods rather than consumers' goods; or, more accurately, he brings about the production of future income (products and services) instead of enjoying present income. The loan is usually made in purchasing power, that is, in money or, more frequently, in deposit money. And the supply of savings that is seeking investment at any time is the supply of purchasing power offered for investment in capital goods and for consumption loans.

The agencies through which savings are directed to the loan market are numerous. Banks, trust companies, mortgage companies, bond houses, investment trusts, and life-insurance companies are in the mar-

ket offering the accumulations of the savers to entrepreneurs and consumers who wish to borrow. The supply of these savings is met by the demands of the entrepreneurs and consumers who wish to make loans. And at any time the rate of interest is determined by the interplay of the forces of demand and supply. The demand of the borrowers is determined momentarily by the condition of trade, by speculative activity in the community, and by all the other forces that bring about opportunities for buying and selling at a profit.)

The demands of the various borrowers are not homogeneous. Merchants and manufacturers make use of a great amount of short-time credit to finance purchases of stocks and of raw materials. They also desire long-time credit to buy permanent equipment. For the former they turn to the banks; for the latter they turn to private investors and to trust companies and other financial agencies that make a business of acting in the role of go-between for savers and entrepreneurs. There is a more or less clear distinction between the markets for short-time and long-time loans. Moreover, investments for a longer period of time are distinguished as to security (first-mortgage bonds versus second-mortgage bonds, etc.) and as to the length of time for which the loan contract runs. There is a distinction also between the market for loans to large corporations of established reputation and loans to smaller concerns that are equally sound but not so well known. Nor are these more than samples of the various types of demand for loans in the modern loan market. We find that instead of one kind of demand for the purchasing power of the saver there are many. Furthermore, as has been said, savers are not indifferent to the kind of investment they make. So, instead of a single loan market, there are many, and in each of them there is at any time a supply of savings seeking borrowers and a demand for savings seeking lenders. The market rates of interest are determined, like other market prices, by the action of demand and supply.

INTEREST RATES AND THE SUPPLY OF MONEY

One of the early explanations of interest ascribed fluctuations in the rate to variations in the supply of money. Since loans are usually contracted in terms of money, it perhaps was natural that men first attributed these fluctuations to the increase or decrease in its supply. In much more subtle forms this same concept persists at the present time.

One special argument for the money theory of interest attaches particular importance to the supply of bank credit. It is well known that when the reserve of a bank increases, its loans also may increase if the banker sees fit to expand them. And the loans may be increased more rapidly than the reserve, since only a fraction of the deposits resulting from loans needs to be covered by reserve. Here, it is said, is a case where the supply of loanable funds depends on the supply of money or, at least, on the supply of substitute money, bank credit. If we examine this statement we must agree with the assertion that an augmentation of reserves will probably result in larger loans at lower contract rates of interest. The banker having the opportunity to do so will generally seek to expand his earning assets, loans, and discounts. The market rate of interest on short-time loans therefore will be affected by an increase of bank reserves. But this effect will not persist. In the first place, the borrowers whom the reduction in rates induces to borrow money will now compete for the physical factors of production and force up their prices. This in turn will necessitate larger loans to finance the same physical volume of trade, and in the end the rise in prices will compensate for and offset the added reserves. In fact, the rising price level usually raises the rate of interest on bank loans. Entrepreneurs gain during periods of rising price levels and therefore are eager to expand their businesses. To do this they need larger loans; and as long as their profits continue to expand, all but the wariest of them will continue to borrow. The augmented demand resulting from rising prices, which result, in turn, from an increase in the supply of money, drives the rate on bank loans temporarily higher than it would have been had no increase in the quantity of money taken place.

But the rise in the price level affects savings. It does so because the increase in bank credit arms the entrepreneurs with more purchasing power and at the same time, through the effect it has upon prices, deprives those who receive fixed incomes of a part of the purchasing power which they formerly enjoyed. The entrepreneurs use the increase in purchasing power that comes to them to increase capital goods at the same time that the power of the fixed-income groups to buy consumers' goods is reduced, and the result is an addition to the capital of the community somewhat in excess of the addition that would have been made if the level of prices had remained constant. But the effect of this process upon the accumulation of capital in the community must not be overstated. Some of the receivers of fixed incomes, notably the well-to-do bondholders and landlords and the better-paid

professional classes, are savers. When the rising price level reduces the purchasing power of the money incomes of these groups it is probable that they reduce both their savings and their consumption. The amount by which they reduce their savings is an obvious offset to the accumulations by the entrepreneurs to whom the banks have extended credit.

If there is a net addition to the capital of the country, the marginal productivity of capital generally will tend to decline, and the demand price for loans also will tend to fall, for a time, at least, because entrepreneurs will be unable to earn enough on the added increments of capital to enable them to pay the old rates of interest. The enforced savings thus produce a decline in the market rates of interest. The decline in these rates will set in motion other forces, however, which will tend to bring about a rise. Many savers will find their incomes reduced, and some may refuse to save and turn to consumption because of the decline in rates.

But, in the long run, interest rates must seek the point where the demand for savings, which is determined by the net marginal productivity of capital, and the supply of capital, which is determined by the marginal cost of savings, are in stable equilibrium. The values of fixed-capital goods, of course, will adjust themselves to changes in the rate of interest. The bank rate, although its fluctuations are governed by many other forces, in the long run must agree with the rate of interest that will equate the demand for and the supply of savings. If the bank rate is temporarily below this rate of interest, loans will be expanded, the price level will rise, and the bank rate also will rise. If the bank rate is above this rate, loans will be contracted, prices will fall, and the bank rate will fall. Thus the bank rate is only temporarily controlled by the volume of money and the reserves of the banks. In the last analysis, the bank rate must seek the level of the rate of interest that capital can earn in the productive process.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Interest was once explained and justified on the ground that the lender had sacrificed the opportunity to use in his own business the funds he lent. Is this a valid explanation?

2. An increase in the productivity of capital goods, say by an invention, would increase the demand for savings. Why? Would it also increase the rate of interest for a time? Would it cause the normal rate of interest to rise or to fall?

3. The banks of a certain city for some time have made use of advertisements to induce people to deposit in savings accounts. Do you think this type of advertisement is likely to prove effective? If it does, will it affect the sales of merchandise in the city?

4. Suppose the productivity of a large area of land were increased by the building of a railroad. How would the rent of the land be affected? Would interest rates in the community rise as a result of this increase in productivity? What would happen to the capitalized value of land rent?

5. How can a pawnshop proprietor exact several times the contract rate of interest that is charged by a bank? Does he actually receive higher interest?

6. "People do not save to obtain interest but to accomplish certain definite objectives—to build houses, to educate children, etc." Is this statement correct? Is it partly correct?

7. In new countries the rate of interest is usually higher than in old countries. How do you explain this? Is it because capital is more productive in new countries? because the supply of savings is less? because time preference is higher?

8. Is a demand for savings indicated (*a*) by a demand for dwelling houses? (*b*) by a demand for short-time consumers' goods, such as clothing? Explain.

9. Suppose that the rate of interest should increase greatly. How would that affect the value of land? of bonds? of automobiles?

10. Great wars are usually accompanied by rising interest rates. This is owing in part to greater risk in connection with all investments. Is this the only explanation?

11. How would the rate of time preference be affected by (1) a steady growth of the national income? (2) extravagance in consumption? (3) old-age pensions paid by the government?

12. Suppose the interest rate should fall to 2 per cent. How would this affect the rentals that tenants pay for houses?

CHAPTER XXVII · Profits

BUSINESS PROFITS AND ECONOMIC PROFITS

The term "profits," as it is sometimes employed in accountancy and generally in business, includes a number of distributive shares. It is a name given to the net earnings of the proprietorship after all necessary payments connected with the business have been made to outsiders and after expenses for depreciation have been deducted. The following statement is illustrative of the derivation of net profits, in a business sense, for a small manufacturer :

Gross receipts		\$155,000	
Deductions from gross receipts			
1. Cost of materials	\$30,000		
2. Labor	40,000		
3. Other manufacturing expenses	5,000		
4. Cost of selling	10,000		
5. Depreciation on plant	8,000		
6. Taxes	6,000		
7. Interest	11,000		110,000
Net income or business profits			\$45,000

The business profits that remain after these deductions from gross receipts have been made are clearly a composite of several economic distributive shares. Since the manufacturer owns the plant, they include interest on capital and rent for the site on which the plant stands ; they include also a portion of the wages of the owner of the business. This composite return, the net gains of the proprietorship, we shall call *business profits*. But the composite return called business profits must be further broken down into the functional shares of which it is made up in order to complete the analysis of distribution.

We shall designate as *economic profits*, or *pure profits*, the gains of the entrepreneur over and above all explicit expenses of conducting the business and all imputed costs, which include the competitive interest on the proprietors' investment, the competitive rent on their land, and the wages they could earn as salaried managers. Stated in another way, economic profits are the excess of selling price of products over and

above the *normal* cost of such products. Recast to conform to our definition of pure profits, this statement stands as follows:

Gross receipts		\$155,000
A. Deductions for payments to others		
1. Cost of materials: wages, rent, interest, and possibly profits	\$30,000	
2. Labor: all wages	40,000	
3. Other expenses: wages, rent, interest, and possibly profits	5,000	
4. Cost of selling: wages, rent, interest, and possibly profits	10,000	
5. Depreciation: wages, rent, interest, and possibly profits	8,000	
6. Taxes: wages, rent, interest, and possibly profits	6,000	
7. Interest: chiefly pure interest, plus some wages, and possibly profits	11,000	
Total	<u>\$110,000</u>	
B. Deductions for imputed costs		
1. Wages	\$10,000	
2. Interest	25,000	
3. Rent	<u>5,000</u>	
Total	<u>\$40,000</u>	
Grand total of all deductions		<u>\$150,000</u>
C. Pure profits		\$5,000

In this account pure profits are much smaller than business profits in the preceding account. This difference is due to the deduction of imputed wages, interest, and rent. But the difference is more fundamental than mere procedure. Business profits do not vary entirely with the efficiency, competitive position, or luck of the concern, whereas pure profits do. Two businesses might show different amounts of business profits simply because the investment of the proprietors was larger in one case than in the other; but when allowance has been made for the imputed distributive shares, no discrepancy can arise from the fact that one concern owns its plant and equipment free from all encumbrance, whereas the other has borrowed a part of the money necessary to purchase these properties. The existence of a pure profit year after year is almost conclusive proof that the concern receiving it either has able entrepreneurs at its head, is engaged in a highly dynamic industry, is working in a highly dynamic environment, or possesses some degree of monopoly power.

PURE PROFITS AS A SHARE IN DISTRIBUTION

The three distributive shares, wages, rent, and interest, accrue to the owners of the three factors of production to which they are ascribed, because of the productive contribution of these factors. If the fourth

factor, entrepreneurship, is a productive agent, it likewise must share in the product of industry, and we must discover for it a marginal contribution to that product. This factor, however, is not entirely comparable to the other three. In Chapter IV the functions of entrepreneurship were briefly described. There it was explained that these functions are always exercised by human beings, but that the work of the men who perform these functions is not the same in a dynamic world as the work of even the most highly paid routine worker. Nor are the qualifications identical for the two kinds of work.

Alfred Marshall characterized the work of a successful independent manager as that of informing himself concerning the markets, materials, and productive processes of his business; of forecasting demands and costs; of meeting new situations with quick decisions and prompt action; and of selecting good subordinates and inspiring them with the spirit of co-operation.¹ The Austrian economist Wieser also has stressed leadership as the distinctive characteristic of the independent manager.² Other economists, emphasizing the work of organization, have found the distinctive productive function of the entrepreneur in the minimizing of losses by careful combination of the agents of production and by the direction of these agents to their most productive uses.

This characterization differentiates the function of the entrepreneur from that of labor of the ordinary routine sorts, but obviously it does not establish a clear *distinction of kind* between entrepreneurship and the highest sort of direction and control by salaried employees. Many plant managers, railroad executives, and managers of large mercantile concerns must have the characteristics and do the sort of work that has been described; yet they receive wages, or salaries, and cannot, so far as the theory of distribution is concerned, be said to exercise a productive function wholly different in kind from that of ordinary labor. If we read the biographies of successful entrepreneurs we find that many of them began as salaried workers in not very responsible positions, later were promoted to managerial or executive jobs, and finally became owners and responsible heads of their own businesses. Nowhere is there a distinct break between the work they did in the subordinate positions and the work they were called upon to do in the higher positions. The differences are all of degree and not of kind.

Land, labor, and capital are distinguished from each other by (1) the differences in the parts they play in production and by (2) differences

¹ Alfred Marshall, *Principles of Economics* (Eighth Edition), pp. 297-298.

² F. von Wieser, *Social Economics*, p. 324.

in the conditions governing their supplies.' The functions of land and labor are especially unlike; the former is the passive agent in production, whereas labor is the active agent. Land furnishes the materials of production; labor moves them about, reforms and recombines them. The supply of land is dependent on unalterable natural conditions, discovery and changes in technique excepted; labor, in a technical sense, is a wasting, replaceable agent. The supply of land is unrelated to cost of production, whereas the supply of labor depends, in part at least, upon its costs of production, both nominal and real. Capital, from the physical point of view, is a product of the other two factors, but its supply is limited by time preference, which is a sufficient distinction to warrant its being set up as a separate factor in production. Entrepreneurship cannot be so clearly differentiated. Its supply is controlled by the same forces that control the supply of labor in general. There are differences between the conditions that make for a large or a small supply of entrepreneurs and other laborers, but the conditions affecting the supply of the former and the supply of professional workers and salaried managers occupying important positions are substantially the same. There is no striking difference between the work of the self-employed business man, whom we must call an "entrepreneur" or else abandon the term, and the work of the most responsible managers.

Here it is necessary to examine controversial matters. It is sometimes said that the real difference between the position of the self-employed business man and the salaried manager is the fact that the former must take sole responsibility for the entire business, whereas the latter is responsible only for the proper conduct of the office he holds, and that if he fails, the burden of his failure will fall on the shoulders of the real entrepreneur. But even this distinction disappears almost entirely when the corporate form of organization supersedes the individual form. If the manager and the directors make a serious mistake, the stockholders and bondholders will bear the resulting business loss. Of course, both manager and directors may find it difficult to obtain other positions and may suffer a loss of personal earnings (wages). But this is not a business risk; it is a risk to personal earnings that impinges on all workers. There is no escape from the conclusion that the entrepreneur is a laborer. But it is said that there is a difference of degree between the work of the hired manager and that of the entrepreneur. Common labor is only responsible for routine work; skilled workmen must perform more difficult tasks; the salaried manager is expected to know routine work and to be able to judge conditions that

cannot be described and provided for in advance ; and the entrepreneur is required, if he is to be successful, to meet still more uncertain conditions outside the business and to control men and materials within the business under situations that are not easily dealt with by routine methods. The difference between the work of the entrepreneur and that of the highest type of salaried official may be very great, or it may be slight.

PROFITS FROM THE EARNINGS OF RARE NATURAL ABILITY

In business, as in the professions, men of exceptional ability make very large incomes. The lawyer whose ability to analyze legal problems or to plead cases excels the ability of all but a few of his competitors receives fees greatly in excess of the average practitioner. As in competitive sports a very slight margin of superiority enables a competitor to capture most of the prizes, so, too, the entrepreneur whose business ability exceeds that of his typical competitors by a narrow margin will receive a labor income greatly in excess of that made by a majority of those in his trade. His marginal productivity is very great, because under his direction the other three factors of production are more productive than they are under the management of men of ordinary ability. He is responsible for these earnings, and they would disappear if he were withdrawn from the combination and a manager of inferior ability put into his place.

It may be asked, however, how this sort of profit differs from wages. For the most part, it differs not at all. Yet there is one important distinction between the earnings of the self-employed business man and the earnings of the salaried manager. When the latter is hired, the owner of the business agrees to pay a stipulated salary for a given length of time, regardless of the success or the lack of success of his management. The agreement will not be continued, of course, if the manager is unsuccessful ; but there is always the risk that the manager will not be successful, and in that case the immediate loss will have to be borne by the owners of the business. Here the emphasis is on *risk*, which is confined to an environment of change and dynamics.

The earnings of executives who possess unusual capacity for leadership in business will always be high, whether they are employees or are self-employed. Often they do not remain in the ranks of the employee class. Not only do their energy and initiative drive them into independent undertakings, but the recognition of their abilities by others enables them to obtain capital so easily that they are seldom allowed to

remain in salaried positions. In large-scale industry many retain the status of employee but at the same time exercise dominant control of the organizations for which they work ; still others retain salaried positions and at the same time conduct independent businesses in other lines.

But the profits of business enterprise do not arise entirely, under present-day conditions, from the earnings of exceptional ability to manage business under difficult conditions. Stockholders in corporations often receive from their investments a return greatly in excess of the going rate of interest, and the same is true of owners of small businesses who are not conspicuously more adept in management than thousands of other men who receive only the ordinary rate of return on their investments in capital. For the most part these extraordinary gains arise from two sources, both interferences with free competition. We shall turn first to the effect of risk and uncertainty on the returns to the factors of production.

UNCERTAINTY IN ECONOMIC LIFE

In a changeful, dynamic society uncertainty is always present. Every person who engages in any economic activity must encounter situations in which the effects of action, of decisions, and of policies cannot be definitely forecast. The laborer takes the chance of bodily injury, of unemployment and abnormally low wages because of an oversupplied labor market, and of adverse changes in the real value of his money wage. The skilled laborer encounters uncertainty in respect to all these, plus the further chance that the trade he has learned at a considerable expense and at the sacrifice of time and energy will be rendered obsolete by an invention or a change in the demand of consumers. Laborers in the salaried class or in the professional and managerial groups run less risk of temporary unemployment and of bodily injury than factory hands and ordinary salespeople ; but in these higher occupations there are also uncertainties of demand for the services of the workers and chances that the type of work for which they have prepared themselves will become obsolete.

The nonhuman factors of production also are subject to uncertainty. Owners of land cannot completely foresee the growth of business in the city where the land is situated, the direction of spread of population and of business-building, the amount of taxation, and many other conditions that affect the rental of their property. Capital invested in any productive enterprise must be invested in the face of numerous uncertainties. Machinery and other capital equipment may be superseded

by improved equipment. The owners may have located the plant in a territory good at the time the plant was built, only to find later that a new plant in a better location has deprived them of their markets. To conclude: all types of economic activity and all the factors of production are subject to uncertainty.

The term "risk" is commonly used to mean the chance of an *adverse uncertainty*. One would not speak of the "risk" that his house would not burn, or that the growth of business in the neighborhood of his land would cause his rentals to increase unexpectedly. Hence we have used the term "uncertainty" to indicate that the unforeseen future events may turn out either more favorably or less favorably than had been anticipated. A risk differs from a speculation in stocks or commodities in that there is no positive gain to offset bodily injury, one of the risks of labor, or to offset loss of capital by fire, windstorm, and the like, which are risks from physical injury. But purely economic uncertainties are of the nature of variations about normal price. The laborer may suffer unemployment, or he may chance to obtain work under conditions where the wage is adjusted on the assumption of a certain amount of unemployment, and yet not be subjected to loss of working time. The owner of capital may find that the demand for its products increases unexpectedly and that his gains far exceed the marginal cost of saving. Landowners receive windfalls from various sources, all of which increase the rental of the land by increasing the demand for it. Purely economic uncertainty gives rise, then, to both unpredictable gains and unpredictable losses.

The fact that uncertainty surrounds the utilization of the factors of production does not mean that each and every owner must take chances of loss or gain in employing them. The owner of buildings may shift the risk of loss by fire to an insurance company; the saver may transfer some of the risk of loss to others by lending to them on mortgage security greater than the amount of the loan. Throughout the economic structure of modern society we find that the taking of risk, the assumption of responsibility for meeting uncertainty, is shifted from groups owning certain factors to other groups who own either different factors or similar factors. But the people who chiefly relieve others of uncertainty are the owners of property—land and capital. Laborers cannot easily guarantee others against loss. Human beings are not vendible assets, and no mortgage can be placed on them to secure other laborers or the owners of land and capital. Examples of the transfer of uncertainty-bearing from one group to another are numerous. When the laborer

receives a stipulated wage, he transfers to his employers the risk that the goods he produces will not sell for anticipated prices. The investor in bonds and in mortgage notes shifts the uncertainty that pertains to the properties in which his money is invested to the stockholder and to the owner of the mortgaged property. But there is no complete shifting. The employer may become bankrupt and fail to pay the stipulated wage; the corporation may fail and the security for its bonds prove inadequate; and the property by which the mortgage note is secured may turn out to have less money value than the amount of the loan.

THE ENTREPRENEUR AS RISK-TAKER

It usually happens in a small business that the responsible head has made a considerable investment of his own capital. This is but natural, since neither laborers nor investors will entrust their services or their property to one who has no assets with which to guarantee, at least partly, the payment of wages and interest. Final authority and the power to guarantee usually go together. The control of a business is not usually entrusted to men unless they can be held responsible for their decisions, and the final test of responsibility for policies and minor decisions is the ability to make good such losses as arise from errors in judgment with valuable property.¹ Thus it comes about that the entrepreneur is both the final authority and the principal risk-taker in the average small business. Naturally this arrangement is satisfactory to the entrepreneur as well as to the investor and the laborer. If a man is asked to pledge his property for the losses that may occur in the conduct of business, he wishes to have the final decision on questions pertaining to the conduct of the business.

In businesses where the giant corporation has become dominant the relation between risk-taking and final authority for business decisions has changed. As we know, the stock of these concerns is often held by thousands of people who own only a few shares each. But a majority of these widely scattered owners of stock never participate directly in the determination of business policies. A small inner group who are large stockholders, or who can count on voting by proxy large blocks of stock, or who have secured control in some other manner, run the affairs of the corporation. Here we find a nearly complete separation of final authority from risk-taking.

¹ There are of course exceptions: trustees, executors of wills, guardians of minors, and the like.

RISK-TAKING AND PROFITS

If one person relieves others of the unavoidable risks of business, he must bear the losses that arise from adverse events; he is also the recipient of gains arising from favorable events. If an undertaking turns out better than was anticipated by a majority of those who had opportunities to invest in it or similar enterprises, pure profits are made; if the outcome is less favorable than was anticipated, net losses are incurred. Now it might be supposed that after a certain type of business had been operating for a generation, the chances of loss and gain would become known, just as it becomes known about how great the probability is that a certain number of houses in each city will burn annually. Something of the sort does take place. Old lines of business are less uncertain than new ventures, which are proverbially speculative. But purely physical phenomena, such as fires and windstorms, against which insurance can be carried, are the results of relatively stable forces. In purely economic affairs there is not the same stability. Even well-established lines of business are subject to the effects of changing demand, inventions, shifts in the location of population, and many other variations that affect costs of production and price. Nothing like the general experience of the fire-insurance companies emerges. Moreover, no insurance actuary can predict whether a given building will burn. Only the general average liability is ascertainable, and even that is not accurately predictable. Whether a given business enterprise will encounter what might be called average conditions is always a matter of uncertainty.

Pure profits arise in part from the good luck of certain enterprises. These are the concerns that encounter fewer than the average adverse conditions. And the profits that come to the proprietors are purely *fortuitous gains*. But since the laborers and the bondholders have been guaranteed wages and interest, the stockholders retain the larger portion of these gains. The investors on mortgage security likewise may secure a part of them, because all contractual interest rates are affected by the risk that the principal will be lost if the mortgagor fails in business. To put the matter concretely, if the investors have lent money to the concern at 5 per cent when government bonds paid only 3 per cent, the extra 2 per cent is chiefly a premium for risk-taking. If the investors find that after a generation no loss has occurred, they may count a part of their receipts from contractual interest as pure profits, and these are also *fortuitous gains*.

In dealing with conditions of uncertainty some men have far greater ability than others; some have far greater opportunities to discover good investments than others; and some have better opportunities for spreading their investments widely, so that, like an insurance company, they will receive the benefit of numerous chances. If the individual has exceptional ability to estimate uncertainty, the extra return on his investments must be attributed to his personal abilities and is therefore a kind of wage. It is substantially the same as the returns that accrue to exceptional ability as an entrepreneur. The gains that an individual receives because he has better opportunities to know the relevant facts about investments than a majority are clearly gains that arise because of the impediments to competition. The same conclusion applies to gains coming from the possession of a large amount of wealth, which enables the owner to spread his investments. Here again pure profits arise out of dynamic conditions.

THE REAL COST OF RISK-TAKING

If the assumption of the greater part of the risks of business were regarded as a hardship by the entrepreneurs, then a special compensation would need to be forthcoming before men would assume these risks. It is often argued that men prefer certainty to uncertainty; that, other things being equal, one would rather have a good investment that paid 5 per cent than a risky one that promised either 10 per cent or nothing. But this generalization is not accurate. Investors are not all of the same temperament. Some actually prefer the chance of great gain, even though it is a very remote one, to an investment that promises a smaller maximum return but a larger probable one. Others can scarcely be induced to take any chance at all. They, too, overlook the actuarial probability. Investors of the cautious type are not usually found among the entrepreneur class unless the inheritance of property forces them into the position. In business they seek salaried positions, and for their investments they demand conservative securities. Entrepreneurs are recruited from those who are willing to take risks.

Whether payment is necessary to induce the bulk of the entrepreneurs to enter business cannot be definitely determined. There seems to be little evidence that anything in addition to the chance that they *may* make large fortunes is necessary. And this conclusion is given more weight by the fact that the independence enjoyed by the entrepreneur is to many men a very great satisfaction. They wish to give orders, not

to receive them, and they wish to manage their own investments. But independence and management are attended by risk, and with control goes responsibility. The real cost of risk-taking does not explain pure profit.

NATURAL LIMITATIONS ON COMPETITION

The theories of value and distribution are usually set forth under the assumption that competition works faultlessly. Yet attention has been called to the fact that interferences exist, and the effects of such obstructions have been explained. Completely free competition is practically never encountered. There is always some immobility of the factors of production, market information is never complete, and human inertia is always present. Nor could these obstacles be entirely removed even if men strove with the best intentions to destroy them. They are natural, a part of the unescapable conditions of human society. It is sometimes said that in a static state competition would be perfect. If there were no change of any sort, men would be able to obtain all the information relevant to their particular tasks, and labor, capital, land, and management would be evenly distributed throughout all industry. With this statement we may agree if we assume also that effects of inertia, impulse, education, and social adaptation also are somehow miraculously abolished. The picture is too nearly perfect. It presents man not as a biologic organism adapting himself to an environment that is both natural and social, but as a machine set to run in grooves. It is a useful device for the first step in economic analysis when we desire, as Alfred Marshall put it, to impound all variables but the one under examination within the pale of "other things being equal"; but it becomes an obstruction to economic thinking if we refuse to recognize its lack of complete reality.

Natural limitations on free competition prevent the free flow of the factors of production from place to place and from business to business. Hence these factors do not always earn their potential normal marginal productivity. Some earn more, and some earn less. The position of the entrepreneur enables him to take some of the gain when by good luck it happens that natural obstacles to competition cause the factors he employs to produce more than the normal rate of return. But he is compelled also to take the lean with the fat; and if the failure of competition works to the disadvantage of the business he controls, he must accept the loss. As we have seen, some of these gains and losses are shared by the laborers, the savers, and the landlords.

Pure profits that arise from these unavoidable limitations on competition, which are not the results of conscious design, are in no sense due to the productive contribution of the entrepreneur. Not only entrepreneurs but also laborers, landlords, and investors participate in surpluses and deficits caused by these obstacles to complete mobility. We cannot find in these fortuitous gains a peculiar share in distribution. They must be looked upon as deviations of the short-run earnings of the factors from their normal or long-run productivity.

GAINS OF MONOPOLY AND OF EXPLOITATION

Any entrepreneur who can buy raw materials, capital goods, labor, and land in the open market at competitive rates and sell their products in a restricted or partly monopolized market can make gains that are above the normal competitive return of these factors. Legal monopolies of all sorts offer opportunities for this kind of gain. So do partial monopolies due to other causes or combinations that restrict the free flow of capital into the fields in which they operate. But these gains are not due to entrepreneurship. Rather, they result from privileges granted by government, from evasions of the law, from environmental obstacles to the mobility of the factors, and sometimes from sheer fraud and intimidation. In the early history of the railroads large shippers sometimes were able to ruin their competitors by forcing the roads to grant them rates below the authorized schedules, which were enforced against their competitors. Today we often hear of "unfair competition" and "unfair methods of getting business." A businessman who induces the employees of his rivals to reveal confidential information, who imitates closely the brands of his competitors, or who deceives consumers may gain by doing so; but the reward he receives is not a payment for productiveness any more than is the gain of the confidence man or the swindler. In every field of productive enterprise there are opportunities for shrewd, unscrupulous, and ruthless men to prey upon their associates in business, upon the consuming public, and upon laborers. Their gains are partly wages and partly profits; the latter we may classify as *exploitative profits*.

In a somewhat different category are the gains from impediments placed by entrepreneurs in the way of the spread of improvements and market information. New processes are carefully guarded to prevent competitors from gaining knowledge of them. The names of profitable patrons, mailing lists, price lists, and the location of desirable markets

are regarded as information not to be made public. Probably the significance of these business secrets has been overestimated. In many lines the producers have found it desirable that market information should be pooled and then disseminated generally to the entire trade. Certainly it is to the advantage of society to have all advances in technique and in business methods made available to all producers. When the entrepreneur can gain by preventing the spread of improvements or of market information, he may earn more on the factors of production than the going rates placed on them in the open market. But we must decline to call such activities productive.

SUMMARY AND CONCLUSION

Business profits, as the term is commonly used, include the total receipts of the business unit less payments to creditors, to laborers, and to others outside the proprietorship. But sometimes profits are composed chiefly of interest, land rent, and the wages of the owner-operators of the business. The surpluses over and above the normal returns to the factors other than entrepreneurship are pure profits. Pure profits can be analyzed into several unlike constituents: (1) Gains due to the uncertainties of economic life, most of which are received by the owners of property. Because the entrepreneur is a property-holder and because of his position in the organization of business, he is in a better position to secure these gains than are laborers or investors who insist on mortgage or other special security. But, in the case of the corporation, risk-taking and the responsibility for final decisions tend to become separated. When they do, the gains from uncertainty go to the stockholders and not to those who control, except to the extent that the latter are also stockholders. (2) Gains from natural or from legally produced artificial obstacles to free competition. (3) Gains from fraud and other illegal and unethical interferences with economic liberty.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Does the farmer perform entrepreneurial functions? the physician? the owner of preferred stock in a corporation? a hired accountant? the superintendent of a plant who owns no stock in the corporation?

2. Is any part of gain from speculation in stocks economic profit? any part of increase in the value of inventories? If not, how should you classify them?

3. Are exploitative profits earned or unearned incomes? Would production be reduced if the government should take them away? Would the prices of

products of the industries in which these profits are found be increased if the government took them?

4. The earnings of especially capable entrepreneurs are sometimes called the "rent of natural ability." In which respects are such earnings like the rent of land? In which respects are they unlike?

5. The entrepreneur has been called the "residual claimant" in distribution. He is said to take the net income that is left after the other factors have been paid their normal shares. Is this statement correct as regards (1) exploitative profits? (2) fortuitous gains? (3) the income from rare natural ability?

6. If you found that a corporation had been earning, on the average, 6 per cent on the original investment of the stockholders, should you say that they had made any pure profit? Suppose that they had earned 20 per cent in one year? 20 per cent on the average?

CHAPTER XXVIII · Trade Unionism and Labor Legislation



THE ORIGIN OF TRADE-UNIONS

Labor organizations are a product of the modern capitalistic economy. They developed first in England, and the following generalizations apply particularly to that country. In the precapitalist period families were to a large extent self-sufficient, producing for themselves the things they needed. In so far as they produced for the market, it was a local and restricted market. The peasant cultivators worked on strips of land which they held under the feudal lord, and they themselves owned the necessary implements and working animals. The handicraftsmen in the towns employed themselves in their own shops and sold their products in the local market. Later came the various inclosure movements, which expelled the cultivators from the soil and made them wageworkers. At the same time the guild system, with its local markets and independent masters, was broken up under the impetus of the growing trade and commerce that sprang from the development of ocean navigation, the discovery of the New World, and the planting of colonies. New industrial towns grew up in which the dispossessed peasants and the former independent masters, who were unable to cope with the new situation, found employment. But it was not a factory system that they entered; it was a domestic, or contract, system. The rising group of merchant capitalists who had contact with distant markets let the work out to be done in the homes of the workpeople. The wage-earners still worked with their own tools, but they worked on materials owned by the merchant capitalists, who sold the products in the distant markets. Sometimes the merchant capitalists employed a limited number of workmen in a "manufactory," where they introduced a considerable amount of division of labor, though the work was done with the aid of simple hand tools. The bulk of the work, however, was done under the domestic, or housework, system. The former independent handicraftsmen were unable to compete in the open market with these merchant capitalists. As a result they were gradually forced to give up their own little

businesses and sell their labor to the merchant capitalists. In this manner the former peasants and independent artisans gradually became a body of permanent wage-earners. Finally came the invention of machinery and the introduction of the factory system. Under this system the wage-earners no longer worked with their own tools in their own shops or homes, but in factories owned by the growing body of manufacturing capitalists. A similar transformation took place in large part in trade, transportation, and other mechanical pursuits.

Thus modern capitalism created an industrial wage-earning class out of a population which had in large part employed themselves. The transformation perhaps can be illustrated most vividly by showing what would take place if American farming should develop from the small-scale, individual-proprietorship type to the large-scale "factory system" type of farming. Suppose that farming could be carried on more economically by corporations owning large tracts of land and employing scores or even hundreds of workmen and expensive machinery. Should such a development take place, our present small-scale, independent farmers would become a body of wage-earners. This is exactly the sort of thing that occurred with respect to a considerable body of peasants and artisans of the precapitalist era. The modern wage-earning class is a product of the Commercial and Industrial Revolutions. Where there had been independent masters owning their own hand tools and selling their own products, there now sprang up in their place wage-earners with nothing to sell but their labor.

In the Middle Ages the handicraftsmen had formed guilds to protect themselves in the sale of their products. Transformed into wage-earners, they now began to form labor organizations to protect themselves in the sale of their labor. This became all the more necessary in the latter half of the eighteenth century and the early nineteenth century because of the breakdown of governmental regulations and the introduction of the laissez-faire regime. Formerly products were sold in a local market controlled by custom and guild or governmental regulations. In the new regime the wage-earners found themselves disposing of their labor in an open, competitive market, a market without any of the dikes and bulwarks of organized effort or governmental protection, a market subject to the inundation of competing labor—woman labor, child labor, prison labor, workhouse labor, and increasing population. The laborer's problem was to protect himself against this inundation, to protect his labor market. Wage-earners became class conscious and began to unite for protection and advancement; and thus trade unionism arose.

HOW TRADE-UNIONS AFFECT WAGES¹

Trade-unions are essentially marketing organizations engaged in selling labor services. We have seen already that wages are determined by the marginal productivity of labor; yet it is well recognized that trade-unions may influence the rate of wages. This fact is in no way inconsistent with the theory of wages presented. At this point the manner in which trade-unions may affect wages will be summarized briefly.

The influence that trade-unions have on wages may be stated under two heads: (1) trade-unions may assist the wage-earner in obtaining from his employer the full value of his labor; (2) trade-unions may alter the value (marginal productivity) of labor.

In the first place, a trade-union may engage exclusively in the orderly selling of the labor supply of a certain trade. Such a union will admit all comers into the union without any restrictions whatever. It seeks to sell the labor of its members for what it is economically worth, namely, the money equivalent of its marginal productivity. Such a union is an informational and bargaining organization. It keeps in touch with all available facts as to the supply and demand for labor, the condition of the market, the prices at which goods are sold, costs of production, the profits of employers, the rates of wages paid in different localities, and the rates of wages in different industries. The union is thus able to bargain intelligently with the employer. Many an ignorant immigrant laborer has sold his labor below the marginal-productivity wage because he knew nothing about alternative opportunities. He did not know what he could get elsewhere. Only when the union came in and informed him what his labor was worth did he obtain the full value of the marginal product of his labor. One of the important functions of trade-unions therefore is to assist workmen in getting wages commensurate with the value of their labor. To do so it is necessary not only to supply information about alternative opportunities, but also to strengthen the reserve or waiting power of the laborers so that they shall not be obliged to take whatever wage is offered. The funds of the union and its power to assess working members constitute such a reserve and strengthen the withholding power of the wage-earners.

In the second place, trade-unions may affect materially the market value or, in more technical economic language, the marginal productivity of labor. The marginal productivity of labor may be affected

¹ See Alvin H. Hansen, "The Economics of Unionism," *Journal of Political Economy*, August, 1922.

in one of three ways: (1) by changing the personal efficiency of labor, (2) by changing the supply of labor relative to the other factors of production, and (3) by changing the efficiency of the other factors of production.

Employers as a rule contend that trade-unions have a tendency to lower the personal efficiency of labor by imposing restrictions on output. Trade-unionists usually insist that they do not attempt to restrict output, and that such restrictions as do take place are common to both union and nonunion workers. Into this controversy we shall not enter except to say that in Great Britain, at any rate, it appears that union regulations have had the effect of restricting output. During the war, however, British unions agreed to abandon temporarily their restrictive regulations in the interest of larger national productivity. In the United States there are indications that some unions at any rate are coming to realize the importance of maintaining production standards.

The effect of trade-unions on wages is more apparent, however, in other directions. Trade-unions always have sought to limit the total supply of labor. Thus they have opposed child labor, woman labor, prison labor, and immigrant labor. They have had a large influence over our immigration policy. The Knights of Labor, an early labor organization, was largely responsible for the passage of the law against contract labor, which prohibited an employer from advancing the transportation fare to a prospective immigrant on condition that he work it out after coming to this country. Labor organizations took the lead in the agitation for the exclusion of Chinese labor. The American Federation of Labor has been in the forefront in the long struggle to secure legislation restricting immigration. More recently the incoming of Mexican labor has been a matter of considerable concern to the trade-union movement. Likewise the trade-unions in the North are opposed to the northward migration of negro labor. Moreover, the labor movement in all countries is opposed to measures which aim to increase the birth rate. It is said that in the period of the Industrial Revolution in England there were weaving districts in which allowances, over and above the regular wages, were made to employees who had more than two children under ten years of age. After the war large-family leagues were organized in France, and numerous funds were set up by employers' associations in several countries in western Europe which paid an allowance (in addition to wages) based on the size of the family. A workman is paid first a wage proportional to the value of the work done, and in addition an allowance for his chil-

dren. The allowance is paid out of a central fund, to which each employer contributes in proportion to the workmen employed. Thus it is immaterial to the individual employer whether he employs single men or family men. It is characteristic that the labor movements of these countries were concerned that these funds should be so organized that they would not tend to stimulate the growth of population.

It follows necessarily from the principles of wages, explained elsewhere, that if trade-unions are able to restrict the total supply of labor, such action will increase the marginal productivity of labor and hence the general level of wages.

But although trade-unions are concerned about the supply of labor as a whole, each union is particularly concerned about the supply of labor in its own special trade or craft. The interests of each labor group are affected, it is true, by the general labor market. A low general wage level is a constant menace to a special high-wage group, since there is always a tendency to spill over from the flooded market to the more favorably situated one. Nevertheless, it is quite possible for a special group to lift its wages far above the general level. Such high wages are to be explained chiefly on the ground of a relative scarcity of labor due to the special skill and training required; but a trade-union may increase this relative scarcity in its own specialized labor market through such devices as the closed shop, high initiation fees, limitation of apprentices, and a minimum wage.

The union closed shop means a shop in which only union people can be employed. If this is accompanied by a closed union (a union which refuses to admit new members or at least limits rigidly the number of members) we have a monopoly control over the supply of labor. Here we have, then, monopoly wages analogous to monopoly price. This does not mean, however, that the wages are higher than the marginal productivity of this labor group. Rather it means that this labor group has artificially limited the supply of labor in this line of work to such an extent that the marginal productivity is exceptionally high.

But many unions do not limit their membership; on the contrary they employ numerous organizers who are constantly seeking to enroll new members. They cannot be charged with being monopolistic in the usual sense of that term. Nevertheless, even though the doors of the union are open to all comers, it may still be that entrance can be gained only after paying an exorbitantly high initiation fee. This policy frequently has been practiced by certain unions, though as a rule the entrance fee is not unreasonably high. Or it may be that an unnecessarily

long term of apprenticeship is required. Such a policy would limit decidedly the supply of labor in the field. As a rule, however, apprenticeship regulations are not of great importance in modern trade-unions.

All trade-unions endeavor, if possible, to establish a minimum-wage scale. If the employers are compelled to obtain all their labor supply through the union, as is the case, for example, in the clothing industry in Chicago, the union, if it finds such a policy advisable, may establish a wage scale considerably above the market rate for that grade of work. As a result, however, a portion of those normally following that trade will be unable to obtain employment. The higher the wage the fewer will be the number employed, since the higher wage scale will necessitate a higher price for the product, and as a result the output that can be sold will be smaller. A union thus may find that while it is able to maintain a high wage level, a large number of its members are out of work, and thus the average annual earnings may be no higher than normal. An abnormally high wage scale thus has the effect of limiting the number actually employed in the trade. Hence the wage rate thus arbitrarily established, although in excess of the marginal productivity of the available supply of labor, is not in excess of the marginal productivity of those actually employed.

The extent to which the number will be reduced if the wage scale is pushed above its natural level depends upon the elasticity of the demand for the labor in question. This, in turn, is affected in part by the elasticity of the demand for the product. If clothing workers receive wages above the natural level, the price of clothing will be higher (constant labor efficiency being assumed) and fewer clothes will be sold. Abnormally high building costs reduce the volume of building. In some lines the effect on the volume of employment would be very slight; in others it would be very great.

In the third place, trade-unions may affect indirectly the efficiency of the other factors of production, and indeed often have done so. This is notably true in the case of the management factor. The pressure for high wages exerted by the trade-union is a constant goad to managerial efficiency. The employer is forced to buckle down to the job of making industry efficient enough to carry the load of high wages. E. A. Filene says that if employers are pressed to pay high wages it will help them "to get away from the habit of thinking about lower wages as the only remedy for high cost of production. The big thing in business and industry that makes for high costs is waste—waste growing out of inefficient labor." "In America," says John A. Hobson, "it is

generally recognized that the high wages obtainable by all skilled and much unskilled labor in industry operated as the potent stimulus to the great economies in mass production, mass transport, and larger use of machines and mechanical power which distinguish their industrial system."

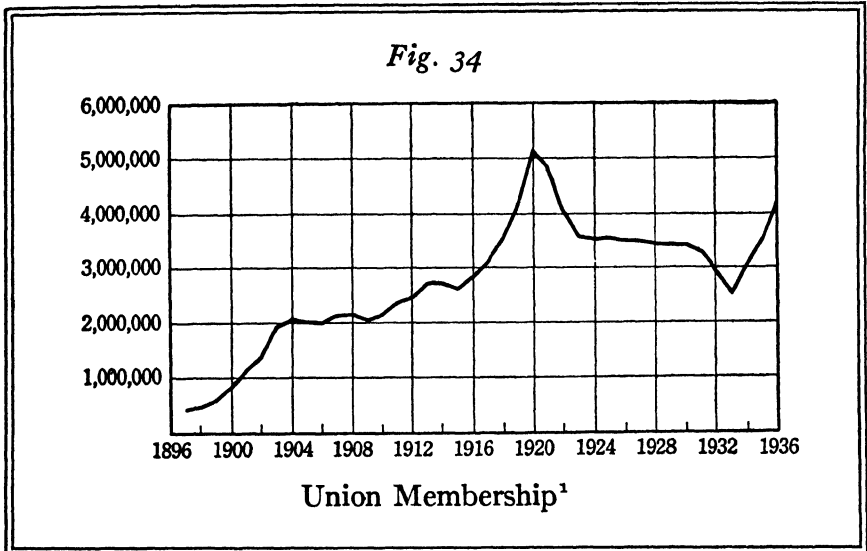
THE GROWTH OF UNIONISM IN THE UNITED STATES

The first nation-wide federation of labor in this country was the National Trades' Union, organized in 1834. This organization sprang up as a result of the upheaval in prices in the thirties and disappeared after the panic of 1837. It was a federation of city centrals, which in turn were composed of local unions in the different trades. The next national federation of labor was the National Labor Union, organized in 1866. During the upheaval in prices incident to the Civil War some thirty national trade-unions—unions of locals whose members follow the same trade, for example, the National Typographical Union—had been formed along with a large number of city federations. The National Labor Union was a federation of national trade-unions and city federations. This organization disappeared in a few years. During the eighties there sprang into prominence an organization known as the Knights of Labor. It became involved in a number of widespread railroad strikes which attracted country-wide attention and brought sharply to the foreground the existence of a labor problem. In the middle eighties a horde of unskilled laborers were inducted into the organization, but after 1886 the membership rapidly declined.

The Knights of Labor, which was a general labor union, soon came into conflict with the American Federation of Labor, an organization built on an entirely different plan. The American Federation of Labor is essentially a federation of national trade-unions. Each of these national unions has complete control over its own strike and benefit funds, engages in collective bargaining, and generally regulates its own policies with reference to dues, fees, apprentices, minimum wage, etc. The American Federation of Labor guarantees that each affiliated national union shall respect the jurisdiction of the other affiliated unions so that there will be no overlapping of work or conflict over members. It employs organizers who help the national unions in their organization work, though the main work of this sort is carried on by each national. Moreover, the federation looks after the legislative interests of the labor movement. It seeks to prevent legislation at Washington which is un-

favorable to labor and exerts its influence to secure the passage of bills that further the interests of the labor group.

The American Federation of Labor was organized in the eighties, but it did not become conspicuous until the period from 1900 to 1904, when its membership increased sixfold. Another enormous jump in membership came in 1916-1920. A number of strong national unions are not affiliated with the American Federation of Labor. Among these may be mentioned the four railroad Brotherhoods—engineers, conductors,



firemen, and trainmen—and the Amalgamated Clothing Workers of America. Nevertheless about four fifths of the total union membership was formerly affiliated with the American Federation of Labor.

A recent development is the formation of the Committee for Industrial Organization, headed by John L. Lewis. The leaders of this movement favor a plan of organization which would include all the workers in one industry in a single organization. On the other hand, the leaders of the American Federation of Labor favor an industrial union of the mass-production workers, but the inclusion of such craftsmen as are employed—carpenters, electricians, machinists, etc.—in their respective craft unions. The Committee for Industrial Organization (C.I.O.) initiated in 1937 an intensive campaign to unionize the steel, automobile, plate-glass, and other mass-production industries.

¹Leo Wolman, *Growth of American Trade Unions*.

It has frequently been asserted that less than 10 per cent of the wage-earners in the United States are organized. This number is an underestimate based on a misconception of the census figures on occupations. Yet it is not difficult to see how this conclusion was arrived at. There are perhaps 5,000,000 organized wage-earners in this country, and there are some 50,000,000 gainful workers. Therefore about 10 per cent of the gainful workers are organized into trade-unions. But the term "gainful" includes businessmen, professional classes, farmers, and farm laborers. If we subtract these groups we find that there are about 30,000,000 urban wage-earners, including women and minors. Therefore over 15 per cent of the urban manual and white-collar wage and salary workers are organized at the present time. This proportion represents considerably less than that reached by unionism in England a generation ago.

Unionism in the United States was formerly concentrated to a considerable extent in certain strongholds, some of which were of unusual importance. There were seven fields in which unionism was particularly strong: transportation, mining, printing and publishing, clothing, building, public service, and theaters. In the great fields of manufacturing and trade, unionism was until recently relatively unimportant. But a number of the union strongholds constituted strategic points in industry. This was notably true of the transportation field.

Of all organized labor 25 per cent was in the field of transportation, 25 per cent in building, 10 per cent in mining, 8 per cent in public service, 7 per cent in clothing, 5 per cent in theaters, and 5 per cent in printing and publishing. About 85 per cent of all organized labor was to be found in these seven fields.

In these seven fields unionism had secured a firm foothold. But the struggle for existence is not over. A large percentage of the bituminous-coal mines are nonunion. The operating railroad men continue to hold firmly the strong position long since attained. The railroad-shop crafts, which rose to great prominence during the period of government operation, subsequently suffered a severe setback, from which they have recently been recovering. Unionism in the building trades has been from time to time in ill repute with the public; it lost ground in the post-war open-shop drive, but is regaining strength.

The great steel and automobile industries, until recently almost completely unorganized, have at last yielded in part to gigantic efforts at mass organization of the workers. In the meat-packing industry, despite the great power which the unions, bolstered up by the government, attained during the war, the membership dropped close to the prewar level

when this prop was withdrawn. Longshoremen, seamen, boilermakers, blacksmiths, machinists, and textile workers, after enjoying an enormous rise in membership during the war period, suffered subsequently a terrific decline compared with the high point then reached. Later, under the stimulus of industrial revival and governmental support for collective bargaining, unionism is now again making headway.

EMPLOYERS' ASSOCIATIONS

Confronting organized labor are organized employers. Some employers' associations recognize and deal with unions; others seek to destroy, or at any rate to weaken, unions. Among employers' associations that have dealt with unions may be mentioned the Stove Founders' National Defense Association, the American Newspaper Publishers' Association, the United Typothetae (closed-shop branch), the Building Trades Employers' Association of New York City and similar organizations in other cities, the Illinois Coal Operators' Association, and the various railroad systems.

Many other employers' associations might be mentioned. The National Council for Industrial Defense was formed to oppose labor legislation. The League for Industrial Rights is especially interested in disseminating information regarding court decisions that deal with certain activities of labor unions, such as the boycott and the sympathetic strike. Of a quite different character is the National Industrial Conference Board, an employers' association interested primarily in research looking toward a more satisfactory adjustment of labor and other problems.

TRADE-UNION DEMANDS

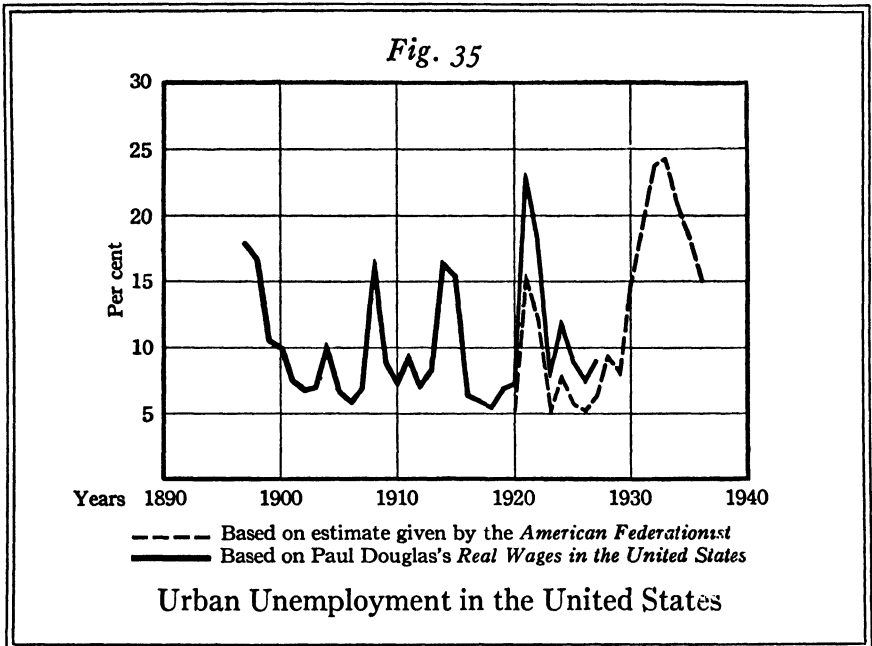
There are those who profess to believe that unions were once necessary but are no longer so. Unions, it is said, served a useful purpose in the old days, when employers exploited their labor and when wages were miserably low and working conditions intolerable. But unions aim at improvement regardless of how high the standard may already be. In fact, unions did not arise among the lowest-paid laborers, nor among those who suffered the worst conditions. The unions formed in this country a hundred years ago were composed almost entirely of skilled artisans. Their wages were nearly twice as high as the wages of the common laborers. Yet it was the highly paid men that organized. The same was true in England.

Nor has unionism abated with the remarkable rise in real wages that we have witnessed during the last hundred years. It is estimated that real wages have approximately tripled during this period. Both the skilled and the unskilled have participated almost equally in this rise. It is true that the present wage level for the great mass of unskilled wage-earners is an unsatisfactory one, but it is vastly superior to that of fifty or a hundred years ago. Yet even though wages were again doubled and trebled, unionism would not disappear. Labor, not unlike other groups, always wants more and more; and particularly is this true so long as we have the existing great disparity in the distribution of wealth and income.

Labor wants not only better wages but more leisure. Here, again, the last century has witnessed a remarkable improvement. A hundred years ago the length of the working day ranged from twelve to thirteen hours, and fifteen to seventeen hours were not unknown. Even as late as 1850 a factory-investigating committee in Massachusetts found the working day to be from eleven to thirteen hours. By 1890 the ten-hour day had become general. By 1930 the average working week in factories was fifty hours, and by 1937 about forty hours. These data give only the length of the standard, or basic, day; they do not include overtime. The working day thus has been reduced to nearly one half its former length, while in the same time real wages have more than doubled.

But labor wants not only good wages and leisure; it also wants security, particularly security of the job. Unemployment is the most serious menace confronting the modern wage-earner, and in this regard thus far we have made no progress. Geographical specialization, together with the continued widening of the market, has made the problem of adjusting supply and demand increasingly difficult. Moreover, we have scarcely begun to grapple with the problem of business stabilization. The business cycle is one of the most important causes of unemployment. Seasonal unemployment also is a grave evil, but labor can more easily adjust itself to the regularly anticipated unemployment that comes and goes with the seasons than with the unemployment caused by business depression. Not only does this come more unexpectedly, but it is likely to last for a considerable period. Furthermore, the causes of cyclical unemployment are not readily understood by wage-earners. Workmen are in constant fear that the job will not last. The restrictive policies of trade unionism are in large part the product of instinctive attempts to make work for everybody. They are crude and ineffective attempts to solve the unemployment problem.

Modern industry, moreover, gives rise to what perhaps may best be described as structural unemployment. The continuous introduction of machinery and other labor-saving devices frequently reduces the number of wage-earners employed in the industries affected. The displaced workmen must be absorbed into new industries and new types of service, and this requires effort, ingenuity, and time.¹ Moreover, there is constantly going on a shifting of industries from old regions to new regions and from old products to new products. Shifts in demand



and changes in technique account for a considerable volume of unemployment, and the more rapid the change the greater the structural unemployment that ensues. During the decade of the nineteen-twenties there was an actual decline in manufacturing employment despite the increase in population, whereas employment in the service industries increased. The railroads reduced their working forces by about 265,000 from 1919 to 1929; there was a great increase in the number of employees in truck transportation, but the railroad workers could not readily shift to the new type of employment.

So long as these technical changes take place, much unemployment will continue to prevail, even though the business cycle should be en-

¹For a further discussion of this point see pages 401-405 of this book.

tirely eliminated. In fact, stabilization of the cycle might even have the effect of making it more difficult to absorb those who had become unemployed on account of changes in technique. Were there no longer any boom periods, it would be more difficult to start new industries and so to absorb the displaced workmen. So long as we have a dynamic society with continual change and progress, a considerable amount of unemployment will necessarily prevail. Unemployment can be minimized, but it cannot be entirely eliminated. Genuine security against unemployment, therefore, makes necessary some form of unemployment insurance. This we shall consider briefly in a later section.

THE METHODS OF UNIONISM

The chief methods by which labor organizations exert pressure to achieve their ends are (1) the strike, (2) the boycott, and (3) political influence. The pressure is in proportion to the strategic importance of the strikers as producers, of the boycotters as consumers, and of the political force of a body of voters that will steadily act together.

The strike weapon appears, if we take a long-run view of the matter, to be increasing in importance. The percentage of industrial wage-earners who went on strike each year from 1886 to 1905 was about 60 per cent higher than obtained for the period 1881-1885; from 1915 to 1921 the percentage was more than four times greater than in the period 1881-1885. During the nineteen-twenties, however, and later during the depression, strike movements greatly declined. From 1922 to 1926 the percentage fell to about twice that of 1881-1885, from 1929 to 1931 to only two thirds as high as in the early eighteen-eighties, from 1932 to 1936 to less than one third that level, but it increased greatly in 1937. Strikes are becoming more and more national in scope. Railroad strikes and coal strikes affect the industrial life of the entire nation. The boycott, which involves a refusal to buy goods of an "unfair employer," although not so formidable as it was some years ago, is still an important weapon of labor, particularly in local disputes. Politically, the influence of labor is growing. This influence showed itself notably during the Wilson administration, and again in the Roosevelt administration.

COLLECTIVE BARGAINING

The strike and the boycott are the chief weapons used by trade-unions to enforce collective bargaining, which is the fundamental goal of unionism. Employers are coming more and more to accept some form of col-

lective bargaining. There are, however, two radically different types: one is the company-union, or employee-representation, type; the other is the trade-union, or industrial-union, type. In the former the company deals only with men who are actually in its employ. The employees may choose representatives to deal with the company, but these representatives are themselves employees of the company. No interference by any outside organization is tolerated. Such organization as exists is limited to those actually employed by that company. On the other hand, a trade-union organization is composed of all workers following the same craft in all parts of the country and employed by hundreds or even thousands of different employers, while the industrial union is composed of all the workers in a given industry. When a union engages in collective bargaining with the employer, it does so through the officers of the organization. These men devote all their time to the work of the union and are not in the employ of the company. They are therefore freer and more independent in their negotiations with employers. Moreover, they have behind them a greater reserve strength both in funds and in membership. The whole force of the national organization thus can be brought to bear upon a single employer. The company union, on the other hand, has no outside force to rely on.

Employers argue that the company-union type is superior from the standpoint of productive efficiency and desirable industrial relations. Since it includes only the employees of a single company, it serves to establish an esprit de corps, a family relationship in the shop, so necessary in maintaining efficiency and industrial good will. On the other hand, unionists contend that such a form of organization leaves labor dependent on and subservient to the employer.

These arguments relate to the "open-shop" controversy. The public mind is befuddled more or less with regard to this controversy because of confusion with respect to the terms used. Strictly speaking, the open shop means a shop in which neither membership nor nonmembership in a union is a necessary condition for employment. It is a shop in which both union and nonunion men may be employed. A closed shop, properly speaking, may be one in which only union men are employed, or it may be one in which only nonunion men are employed; an open shop is one in which both union and nonunion men may be employed. There are industries which run on the open-shop principle in which nevertheless labor is well organized and in which unions deal collectively with the employers. The railroads are a good example. The employer is free to hire a nonunion man, but he must pay him

the union wage and give union conditions of work; strictly speaking, his business is an open shop.

Popularly, the terms are used in a very different sense. As the public generally understands the terms, the closed shop means a shop in which the union is recognized and bargains with the employer, whereas the open shop means one in which the union is not recognized by the employer. As popularly used, the terms are therefore misleading. Because of this misuse of terms it is difficult for the public to understand that union recognition and collective bargaining are not incompatible with the genuine open shop. Indeed, in order to maintain a real open shop, it is often necessary to secure a union agreement which protects union men in their right to be employed. This confusion of terms has gone so far that employers who insist that their workmen must sign contracts to the effect that they will not join a labor union publicly proclaim that they are running an "open shop."

While there is much talk about the closed shop, the real opposition is for the most part to unionism itself. The great bulk of employers in the United States are still opposed to collective bargaining through trade-unions. In Great Britain, on the other hand, the majority of employers take trade-union collective bargaining as a matter of course. But that attitude is a product of a long period of evolution. During the greater part of the first three quarters of the nineteenth century British employers fought the principle of trade-union bargaining as bitterly as American employers do now; but the growing power of labor organizations forced the employers finally to deal with the unions. No doubt many British employers would be glad to free themselves from the limitations imposed upon them by these agreements, but unions are too firmly entrenched in the British social order. In the same manner, American employers who now recognize and deal with unions do so not because they prefer this system, but merely because there is no practical alternative.

UNIONISM IN THE MODERN INDUSTRIAL ORDER

The increased growth and influence of labor organizations in all modern industrial countries point to the conclusion that they are inevitable concomitants of the modern industrial order. William Howard Taft, former President and later Chief Justice of the Supreme Court, has given clear and concise expression to this view in the following sentences: "Organization of labor has become a recognized institution

in all the civilized countries of the world. It has come to stay; it is full of usefulness and is necessary to the laborer. . . . There is among employers the Bourbon, the man who never learns anything and never forgets anything; the man who says: 'It is my legal right to manage my business as I choose, to pay such wages as I choose, to agree to such terms of employment as I choose, to exclude from my employment union men, because I don't approve of the tenets of the union, and to maintain a family arrangement of my own. I do fairly by my men; I pay them what I think is right, and they will not complain unless some outside union agent interferes. I run a closed nonunion shop and I am happy and propose to continue happy.' This man is far behind in the progress of our social civilization. He lacks breadth of vision extending beyond the confines of his shop. . . . He does not recognize that we have advanced beyond the state in which employers and employees are mere laws unto themselves. . . . But whether we will or not, the group system is here to stay, and every statesman and every man interested in public affairs must recognize that it has to be dealt with as a condition, to be favored in such a way as to minimize its abuses and to increase its utility."¹

There are many employers, however, who are not willing to accept this conclusion. They believe that the labor problem can be handled better without union organizations. They hold that the union brings industrial war rather than industrial peace, and that it is impossible to arrive at any high level of efficiency with labor and management organized in opposing camps. Hence they favor group insurance, welfare work, profit-sharing schemes, sale of company stock to employees, shop committees, and company unions.

There appears to be no way in which wage-earners can make sure of an independent position in their negotiations with employers except through organizations controlled exclusively by themselves. For this reason it seems probable that unionism will continue as an important institution, at any rate in democratic societies. It is true that authoritarian states have rooted out unionism and ruthlessly suppressed strikes. This a democratic country could not do and continue to survive.

Should a fascist type of government eventually be established in the remaining democratic countries, it is likely that this development will come in consequence of a gigantic struggle between organized capital and organized labor, a struggle in which two opposing monopolistic

¹Daniel Bloomfield, *Problems of Labor*, pp. 212-214.

groups refuse to co-operate in a workable compromise, a struggle in which the state, driven by public opinion incited by fear of violence and disaster, will rigorously suppress one side or the other. If this struggle should lead to domination by labor, we should then arrive at a communist fascist state; if it leads to domination by employers, a capitalist fascist state. Self-discipline, restraint of class interests, capacity to work out an orderly solution through the processes of collective bargaining, these continue to be the methods and the hope of democratic societies.

THE LEGAL STATUS OF UNIONISM

From the legal standpoint trade-unions are a recognized part of our social order, but while trade-unions are lawful organizations in the United States, many of the rights claimed by unionists are seriously questioned by the courts. There have been a number of court decisions which severely restrict the *freedom to organize*. Thus the United States Supreme Court held, in *Adair v. United States*¹, that an employer has a constitutional right to discharge a man because he is a union man. This decision thus gives the employer a right to maintain an antiunion closed shop. Again, in *Coppage v. Kansas*² the court held that an employer has a constitutional right to force his employees as a condition of employment to sign contracts that they will not join a labor union while they are in his employ. Moreover, in the *Hitchman Coal and Coke Company* case³ the court held that it is illegal to try in any way to organize employees if they have signed a contract that they will not join a labor organization.

It is sometimes said that trade-unions are not subject to any control in the United States; that they are irresponsible bodies, since they are unincorporated. As a matter of fact, trade-union activities are subject to considerable legal control in the United States, and indeed labor chafes under the present degree of restriction. While strikes for higher wages and shorter hours are regularly considered legal, in the majority of states there are many forms of strike that are held to be illegal, such as the strike for a closed union shop, a sympathetic strike entered upon for the purpose of aiding some other group of workers to gain their ends, and strikes against the use of nonunion materials. The unions are restricted likewise in the case of the boycott, which is an organized refusal to purchase or handle the products of a merchant or manufacturer.

¹ 208 U. S. 161.

² 236 U. S. 1.

³ 245 U. S. 229.

The secondary boycott, that is, a boycott which involves third parties, such as the refusal to buy from retailers who handle the products of the manufacturer against whom the union has a grievance—is held to be illegal in almost all jurisdictions. In many jurisdictions the primary boycott, which does not involve third parties, is held to be legal. Thus in many cities the labor movement advertises through its press and by banners that certain stores are unfair to organized labor, thus seeking to persuade labor sympathizers to withhold their patronage. Such primary boycotts are quite generally upheld by the courts.

Picketing, that is, efforts on the part of strikers to prevent the employer from securing new employees, is generally held to be legal provided it is carried on peacefully and without coercion. Thus, if at each entrance to the plant the strikers have one man carrying a banner which states that a strike is on and urging workers to stay away, such action is generally sustained by the courts. If, however, hundreds of employees are massed in front of the struck plant, such action can scarcely fail to intimidate prospective employees and may amount to coercion. When coercion, intimidation, and violence are in evidence, picketing is generally held to be illegal.

The employer has two main remedies at law if his business suffers unlawful injury at the hands of organized labor. In the first place, he may institute a suit for damages. The most famous suit is probably the one instituted against the Danbury Hatters' Association for injury suffered by the Loewe Manufacturing Company as the result of a nationwide boycott. In this case¹ damages amounting to about \$300,000 were levied against one hundred and seventy-five members of the hatters' union. The members were held liable because they were members of the union whose officers were conducting the boycott.

The decision in the Coronado Coal Company case² goes further. It holds that the funds of the union itself, as well as the property of the individual members, may be seized in payment of damages resulting from the unlawful acts of the union. In the Coronado case, however, no damages were actually assessed against the United Mine Workers of America, since it was found that the national organization had not authorized the strike in which the unlawful acts in question had occurred. Thus the property of the union as well as the property of the individual members may be taken for damages resulting from unlawful acts authorized by the union.

¹Loewe v. Lawlor, 208 U. S. 274.

²268 U. S. 295.

In the second place, the employer may apply to the court for an injunction, which is an order from the court prohibiting unlawful activities that would result in irreparable damage to property. Injunctions seek to prevent acts that would result in loss to the business in question; damage suits seek to recover damages actually sustained.

It is clear that both the strike and the boycott may result in severe business losses even though no physical property is in any way destroyed; but the mere fact that losses are in prospect or are actually sustained is no legal ground in itself for either the issuance of injunction or the sustaining of damage suits. The losses must follow as a result of *unlawful acts*, or the suits for injunction proceedings or damages would not be upheld in the courts. The unions want a broadening of the law so that all forms of strikes and boycotts shall be considered lawful acts. This, however, does not imply that violence or physical injury to persons and property should be condoned by the law; for any person who commits violence to person or property would still be held in the criminal courts. The unions wish, however, to permit wage-earners collectively to withhold their labor and buying power from any person for any purpose whatever, just as an individual may now lawfully do; in other words, they wish trade-unions to be freed from the law of conspiracy. This freedom they have not had in the United States. For example, in the Bedford case¹ the Supreme Court held that the union was guilty of a conspiracy in restraint of trade because its members collectively refused to work on stone quarried by nonunionists. The Supreme Court held, in *Truax v. Corrigan*², that an act freeing unions from the law of conspiracy is unconstitutional. The Norris-La Guardia Act of 1932 sought, however, to limit the scope of injunction proceedings by providing that no Federal court may prohibit workers, singly or in combination, from engaging in a long list of activities which are specifically enumerated.

GOVERNMENT INTERVENTION IN LABOR DISPUTES

As strikes have grown in scope and importance, the government has increasingly put forth efforts to settle labor disputes peacefully, in order to protect the public. Especially has this been true in regard to disputes on the railroads and in the coal mines, industries in which continuity of service is of vital importance to the public. A compulsory-investigation act applicable to the disputes of the railroads was passed in 1888.

¹274 U. S. 37.

²257 U. S. 312.

It remained on the statute books for ten years, but was made use of only once (in the Pullman strike in Chicago in 1894) and then to no effect. In 1898 the Erdman Act was passed, providing for mediation and voluntary arbitration. The act was restricted to men actually engaged in the operation of trains. Upon request from either party to the dispute, the chairman of the Interstate Commerce Commission and the Commissioner of Labor were to attempt to adjust the dispute by mediation. If that failed they were to urge both parties to submit the dispute to arbitration. Up to 1906 the act was made use of only once, and then to no effect. From this time on, however, railroad rates were more rigidly regulated, and hence it became more difficult for the railroads to grant wage increases. From 1906 to 1913 the act was made use of successfully forty-eight times. In no case did either side repudiate an award. In 1913 the Newlands Act was passed. This act provided for the establishment of a Board of Mediation and Conciliation which could act on its own initiative or by request of either party to the dispute. If the board could not mediate the controversy, it was to urge arbitration. Any question arising with respect to the interpretation of an arbitration award was to be referred back to the arbitrators. During the next four years this act was utilized successfully in over seventy cases. But in 1916 arbitration was refused by the unions, and a general railroad strike was imminent. It was averted by the Adamson Act, which granted the basic eight-hour day on the railroads. Late in 1916 President Wilson asked Congress to pass a compulsory-investigation act with compulsory cessation of strike or lockout during the period of the investigation, but no law was passed. Late in 1917 the government assumed control of the railroads. A Railway Wage Board was established, and also Boards of Adjustment which took up all disputes other than those relating to wages. Finally, in 1920, when the roads were returned to private control, a United States Railroad Labor Board was established with power to decide all disputes (including wage disputes) by majority vote. However, no penalties except publicity were provided for failure to comply with the award. The act was in no sense, therefore, a compulsory-arbitration act. Nevertheless the power of public opinion behind the board's award was an important factor in the failure of the shopmen's strike of 1922. In 1926 the Watson-Parker Act was passed, which abolished the Railroad Labor Board and established in its place a Board of Mediation. This act provided for the settlement of disputes by conferences between the interested parties and by voluntary arbitration. The Railway Labor Disputes Act of 1934 continued this general

policy. It provides for a National Railroad Adjustment Board consisting of representatives of the railroads and the employees and designed to settle disputes by conciliation. If this fails, mediation is attempted by the National Mediation Board consisting of three impartial members. Such mediation may result in submitting the dispute to voluntary arbitration.

In the case of the coal-miners, the Federal government in recent years has intervened frequently in an effort to prevent strikes. In 1919 the Department of Justice secured an injunction prohibiting the officials of the United Mine Workers from distributing strike benefits and compelling them to rescind the strike order. Later the President induced both parties to submit the dispute to a board appointed by the President. Moreover, the Secretary of Labor at times has acted as mediator in coal strikes, as well as in industrial disputes, frequently with considerable success.

During the war the National War Labor Board was established by Presidential proclamation. No special legal provisions were made for the enforcement of its awards. The appeal to patriotic sentiment was generally sufficient to obtain compliance with its orders. If this was not sufficient, the war powers of the President, by which he could take over essential industries and impose the penalty of loss of industrial exemption, were resorted to. Only in very few industries was it necessary to resort to such drastic measures. Practically, therefore, the War Labor Board had powers of compulsory arbitration during the war period. The board was discontinued, of course, at the end of the war.

In 1935 the Wagner Act was passed, establishing the National Labor Relations Board. The act seeks to guarantee to labor the right of collective bargaining through representatives of their own choosing. Employers may not dominate or interfere with the formation or administration of any labor organization and may not refuse to bargain collectively with the representatives of their employees. The new board may determine, by secret ballot of the employees or otherwise, what is the proper agency for collective bargaining in any particular case. The board may order employers to cease and desist from unfair practices and may secure enforcement of such orders through the Federal circuit courts.

Only one American state has tried the experiment of compulsory arbitration. In 1920 Kansas established an Industrial Court with powers of compulsory arbitration in all industries engaging in the transportation or manufacture of fuel, food, and clothing, or in the construc-

tion of houses. Men were free to leave any employment; but strikes, picketing, and lockouts were prohibited. The court did not succeed in preventing large strikes, though it served successfully in the settlement of small disputes. In 1923 the Supreme Court of the United States declared unconstitutional that part of the law which pertained to wage-fixing in such industries as meat-packing.¹ In 1925 the Supreme Court held further that hours of labor cannot be fixed in such industries.² The net effect of these decisions was to restrict the operation of the act to those industries long recognized in law as public utilities. In 1925 the Kansas legislature abolished the Industrial Court and transferred such powers as it still retained to the Public Service Commission.

In Canada a compulsory-investigation act with compulsory cessation of strikes or lockouts during the period of investigation was passed in 1907. The award of the investigating board is not compulsory. In so far as investigating boards have been applied for, the act has helped to settle disputes, but in the majority of cases it is violated at the outset by refusal to apply for the establishment of a board before a strike or lockout is declared. Penalties for violation of the act have not been enforced.

In Australia compulsory arbitration obtains in most of the states, as well as in the commonwealth itself. New Zealand also has a system of arbitration. Victoria has minimum-wage boards, which set up standard minimum wages for almost all industries. Strikes for wages above the minimum set by the wage boards are not illegal, and yet this state has been freer from strikes than some of the states with compulsory arbitration. New South Wales especially has had serious labor disturbances, largely because of the turbulent conditions in the coal mines, which are of considerable importance in this state. Compulsory arbitration has not prevented strikes in any of the states, though disputes of an interstate scope in the main have been handled successfully by the commonwealth court. Compulsory arbitration, however, has furnished the machinery for negotiation and conciliation between employers and employees. The tendency is more and more to work through joint conferences of the parties to the dispute. Compulsory conference is made use of if necessary. The system therefore may be properly described as legalized collective bargaining.

In the United States the labor movement is strongly opposed to compulsory arbitration. The American Federation of Labor believes it can best attain its ends through trade-union activities rather than

¹ 262 U. S. 522.

² 267 U. S. 552.

through industrial courts or labor boards. In the past it has been interested in legislation chiefly to secure freedom from the restrictions imposed by law upon the activities of trade-unions. It has been interested also in labor legislation for children and women and in certain special types of legislation for men. More recently it has favored an extension of governmental control. It favored the labor provisions of the NRA under which minimum standards relative to hours and wages were enforceable by law. After the NRA was declared unconstitutional it favored new legislation designed to regulate wages and hours.

MINIMUM-WAGE LEGISLATION

Twenty-two states have passed minimum-wage laws applying to women and minors. These are Utah, Massachusetts, Minnesota, Wisconsin, Washington, California, Colorado, North Dakota, South Dakota, Oregon, Connecticut, Illinois, New Hampshire, New Jersey, New York, Ohio, Kansas, Nevada, Oklahoma, Rhode Island, Arizona, and Arkansas. Minimum-wage boards usually establish the rate for each industry affected. These boards, on finding by investigation that wages are insufficient in a certain industry to maintain the specified standard of living, may establish a subordinate wage board composed of representatives of the employers, the employees, and the public. These boards, on investigation, recommend the establishment of a certain minimum wage for the industry. Then a public hearing is held by the permanent wage board or industrial commission, and finally a scale of wages is ordered to be made effective. In most states this award is compulsory, but in Massachusetts the only penalty is publicity in the newspapers regarding firms which refuse to pay the established wage.

By a four-to-four decision the United States Supreme Court upheld the Oregon wage law in 1917; but in 1923 it declared unconstitutional the minimum-wage law passed by Congress for the District of Columbia.¹ In 1925 the United States Supreme Court refused to sustain the Arizona law, and subsequently a Federal court in Wisconsin held that the law in that state could not be applied to adult females. The Minnesota supreme court, however, held that the law of that state was valid in its application to minors, and in no jurisdiction has any court held to the contrary. The net effect was that these laws could legally be enforced only with respect to children and minors. To meet a part of the

¹ 261 U. S. 525.

specific objections raised by the court, some states provided that the minimum wage should be equal to the productivity of the worker, or to the specified standard of living, whichever was lower. The revised New York statute was declared unconstitutional in March, 1936. Early in 1937 the United States Supreme Court reversed its decision and upheld by a vote of five to four the Washington state minimum-wage law. Thus at last, after many years of judicial disapproval, this legislation was held valid by the highest court.

HOURS OF WORK

Legislation regulating the hours of work is on firmer legal grounds. Most of the important industrial states limit the hours of labor to eight a day for children under sixteen years of age. About half the remaining states have a nine-hour day. Some of the Southern states still permit ten hours of work a day for children. Nearly all states limit the hours of labor for women, many to eight or nine hours a day, but some place the limit at sixty hours a week. The Oregon law, providing for a ten-hour day for all employees in every mill, factory, or manufacturing establishment, was upheld. The hours of labor are limited for employees of local, state, or Federal governments. On the railroads the hours of labor are regulated in the interest of public safety. Because of the hazards and dangers to health in the mining industry, almost all the important mining states limit the hours of labor for miners to eight hours a day. On similar grounds a few states enacted eight-hour laws applying to adult males employed in laundries, electric plants, plaster and cement mills. Ten-hour laws have been applied in certain cases to sawmills, bakeries, brickyards, and textile mills. Hours of labor were regulated under the National Industrial Recovery Act, which was enacted in 1933 but declared unconstitutional in 1935.

SOCIAL INSURANCE

Turning now to social insurance—insurance for wage-earners against accident, sickness, old age, and unemployment—we find legislation in the United States far behind that of European countries. Until 1935 only in the case of industrial accident was the workingman protected by law in the United States. Practically all states, as well as the Federal government, now have legislation providing for compensation for injuries sustained in industrial employment. The expense is borne exclusively by the employer, who generally is compelled to insure either

in a state fund or in a private company. If the employer refuses to come under the compensation act he cannot set up any of the old common-law defenses in a suit for damages, such as (1) contributory negligence of the worker himself, (2) negligence of a fellow employee, or (3) assumption of risk on the part of an employee when he takes the job. The compensation usually amounts to from half the wages to two thirds, with certain upper and lower limits beyond which the compensation cannot go.

In Great Britain the wage-earner is insured not only against accident but also against sickness, unemployment, and old age. In the case of accident the employer alone bears the expense, as in the United States; in the case of sickness and unemployment the expense is borne jointly by the employer, the wage-earner, and the state; in the case of old age joint contributions are made by employers and employees in the contributory insurance scheme under which benefits begin at the age of 65, while noncontributory state pensions are paid to all indigent and needy persons aged 70 or more.

Many employers in the United States have instituted special funds for their own employees, insuring them against sickness, old age, and even, in some cases, against unemployment. A few companies have engaged to guarantee their workers a minimum number of weeks of employment a year, and to pay unemployment benefits if the employment falls below that level. Such plans have usually been modified seriously under the pressure of the adverse conditions in the great depression which began in 1929. The men's clothing industry in Chicago, in conjunction with the Amalgamated Clothing Workers, instituted some years ago an employment-insurance plan, supported jointly by the employers and the union, which provided for the payment of unemployment benefits at the rate of 30 per cent of full-time wages for a period not to exceed seven and one-half weeks per year. During the depression it was not found possible to pay benefits for the maximum period.

SOCIAL SECURITY ACT OF 1935

The more important features of this act provide for (1) Federal aid to state pensions for needy aged persons; (2) a national system of old-age insurance for some 26 million wage-earners; and (3) Federal stimulation and support to state unemployment insurance.

The *old-age pension* provision of the act gives encouragement and support to state legislation. The Federal government will match state

benefit payments dollar for dollar up to \$15 a month, thus providing a total benefit payment of \$30 a month for needy old persons. States of course may make more generous benefit payments, but Federal aid is limited to a maximum of \$15 per month. State laws must provide pensions at age 65 after 1940, until which date age 70 is permissible.

According to the President's Committee on Economic Security there were 7,500,000 persons 65 years of age in 1935. Of these one half were dependent upon others for support. At the beginning of 1937, old-age pensions were being paid to 1,200,000 persons in 41 states. By 1937, 47 states had submitted old-age pension systems for approval by the Social Security Board.

The *national system of old-age insurance*, unlike the old-age pension system, is not based on the principle of need. Under the insurance plan, old-age benefits are paid as a matter of right, regardless of the need. This follows logically from the fact that it is a contributory system. Workers and employers are required under the act to make equal contributions from wages and from pay rolls, 1 per cent from each for the first three years beginning with 1937. Thereafter the rate is increased by 0.5 per cent for each three-year period until by 1949 the contributions will reach a total of 6 per cent, half paid by the workers and half by the employers.

If an employee receives more than \$3000 per annum, contributions are made by himself and his employer on only the first \$3000 of his earnings. The act covers all gainful workers except proprietors, self-employed persons, the professional class, farm laborers, domestics, teachers, governmental and institutional employees, and casual laborers. Altogether, the act covers about 26,000,000 employees.

Benefits are not payable until the year 1942. Benefits vary according to the total wages received (not counting wages in excess of \$3000 per annum) after December 31, 1936, and before age 65. The monthly benefit equals $\frac{1}{2}$ of one per cent of the first \$3000 total wages, $\frac{1}{2}$ of one per cent of the next \$42,000, and $\frac{1}{24}$ of one per cent of total wages received over \$45,000. Thus a worker who had earned \$100 a month for 40 years would be entitled at age 65 to a monthly benefit of \$51.25 for the remainder of his life. The maximum monthly benefit that may be received is \$85. If the worker dies before reaching age 65, the heirs will receive 3.5 per cent of the total wages received by him after December 31, 1936. If he dies after age 65 but before his benefit payments had equaled 3.5 per cent of his total earnings, his heirs will receive the residue.

In time the insurance plan would presumably diminish the necessity for pensions to needy old people, but it cannot be expected to supplant it entirely. The insurance plan does not cover all employees; nor does it cover professional or self-employed persons, many of whom may find themselves in want in old age. Moreover, provision must be made for the unemployables who reach old age. Finally, pensions will be needed in those cases in which the annuity provided by the insurance fund proves inadequate to care for the insured and his dependents.

The rates of contributions to the insurance system at first are higher than necessary on a pay-as-you-go plan, and are calculated to build up a fund approximating \$47,000,000,000 by 1980. This huge reserve arises partly from the fact that no benefits will be paid until 1942, while contributions begin January 1, 1937 and partly from the fact that the first recipients of benefits will receive reduced monthly payments calculated, not on the basis of the normal total earnings of a worker aged 65, but only on the earnings received after 1936. Thus the receipts for many years will exceed the benefit payments. As this fund gradually accumulates, interest will be earned on the investments. These receipts (at an interest rate of 3 per cent) will supplement the 6 per cent tax on wages and pay rolls combined, and, indeed, by 1980 will amount to more than half of the combined receipts from these other sources. Thus in the early years, while the fund is being built up, workers and employers will be paying to the system more than is taken out in benefit payments. But once the young workers of 1936 (the year the plan was started) reach retiring age (about 1980) and the system reaches a point of relative stability, it will be found that the contributions of workers and employers are each year less than the benefits paid out from the fund by the amount of the interest receipts from the investment of the old-age reserve fund. At any rate this will be the case if the estimates of the actuarial experts of the Committee on Social Security prove substantially accurate.

On a pay-as-you-go plan, less than 3 per cent contributions from workers and employers would be needed in the first years. But eventually, as such a plan came into full operation, the rate would have to be raised, or else the government would have to make up the deficit.

The main arguments against the accumulation of a huge reserve are (1) the possible difficulty of adequate investment outlets for so large a sum, (2) the danger that the fund will be dissipated on larger annuities, and (3) the danger that the presence of the fund will encourage wasteful borrowing by the Federal government.

The *unemployment insurance* section of the act provides, not for a nation-wide system, but for Federal encouragement to the enactment of state laws. The act imposes a Federal tax of 1 per cent for the year 1936, 2 per cent for 1937, and 3 per cent per annum thereafter on the pay rolls of employers having eight or more employees. This tax is applicable to employers in all states. If, however, a state passes an unemployment insurance act and finances the same by imposing a pay-roll tax, the employers may deduct from the Federal tax any sums so paid, up to 90 per cent of the Federal tax. Thus if a state imposes a pay-roll tax of 2.7 per cent, an employer will pay 2.7 per cent to the state and 0.3 per cent to the Federal government. If, however, the state imposes a 3 per cent tax, the employer will still be compelled to pay 0.3 to the Federal government, making a total of 3.3 per cent. If a state imposes a 2 per cent tax the employer can deduct only 2 per cent from the Federal 3 per cent tax, leaving a tax of 1 per cent payable to the Federal government.

Most of the 45 states which by 1937 have passed unemployment insurance acts have imposed a 2.7 per cent pay-roll tax. This, as shown above, combined with the residual Federal tax of 0.3 per cent, imposes a total pay-roll tax of 3 per cent upon the employer. But if the state enacts no unemployment insurance law, the employer will still have to pay a 3 per cent Federal tax. Thus the employers in any one state are not placed in an unfavorable competitive situation by the enactment of a state unemployment insurance law. It is the essential purpose of the Federal act to remove the interstate competitive disadvantage which a state with unemployment insurance would suffer, without a Federal act, as against states with no such law.

The Federal act gives wide freedom to each state to enact whatever type of unemployment insurance act it may prefer. There are five main possible types: (1) the simple pooled-reserve plan; (2) the pooled-reserve plan with merit rating; (3) the individual-plant-reserve plan; (4) the individual-plant-reserve plan with some minimum contribution to a central pool; and (5) the guaranteed-employment plan.

The simple pooled-reserve plan provides that all employers must contribute equally to the unemployment fund regardless of their employment record. This plan is usually opposed by employers in industries with relatively stable employment, since it compels good employment risks to help pay unemployment benefits in industries suffering from severe unemployment. This plan, however, offers greater security for the workers.

The pooled-reserve plan with merit rating compels every employer to make his contributions to a central state fund, but his contributions may be reduced if the employment record of his firm shows high stability. However good his record, he must pay at least something, say a minimum of 1 per cent of his pay roll, to the central pool. The Federal act permits him to deduct from the Federal tax the full normal state tax up to 2.7 per cent, even though the particular tax which the state imposes on him is, owing to his good employment record, below the normal state rate. The merit-rating plan seeks to adjust the contributions in accordance with the risk of unemployment. It is therefore fairer to employers whose industries are relatively stable. The protection to the worker is, however, reduced unless the more unstable industries are compelled to contribute in excess of the normal state tax, thereby making the general average rate actually paid 2.7 per cent of the total state pay roll.

The individual-plant-reserve system permits each employer to maintain his own fund. An unemployed worker is paid benefits out of the reserve of his own employer. If this is exhausted, benefits cease. There is no central state fund. Once the plant reserve has reached a certain sum—in the Wisconsin law, 10 per cent of the annual pay roll—no further contributions are required. If states permit individual-reserve funds, the Federal act requires a minimum plant reserve of 7.5 per cent of the last year's pay roll and not less than five times the largest amount of compensation paid out within any one of the three preceding years.

The individual-plant fund with a minimum contribution to a state fund, say of 1 per cent of the pay roll, represents a compromise between individual-plant funds and pooled funds. It furnishes some added protection to the worker beyond the fund accumulated by his own employer.

Under the guaranteed-employment plan, employers guarantee at least 40 weeks of employment a year. Only in the event that they are unable to fulfill this obligation do they pay unemployment benefits. Thus the worker is assured of at least 40 weeks of wages (thirty hours a week) in each year. If this plan is used, employers must build up a reserve at least equal to 7.5 per cent of a full 40-week pay roll.

Forty-five of the states had passed unemployment-insurance laws by 1937. These cover the great bulk of wage-earners. Benefits are not paid during the first weeks of unemployment, the waiting period ranging from 2 to 6 weeks in different states. The duration of the benefits is limited to a specified period, 16 weeks being the most usual. The amount of the benefits paid is most frequently fixed at 50 per cent of the wage received when employed.

ECONOMIC PRINCIPLES OF LABOR LEGISLATION

We have now to consider how labor legislation affects the production and distribution of the national income. We shall consider this under two heads: (1) how labor legislation affects the *personal* efficiency (and therefore the marginal productivity) of the wage-earning class; (2) how labor legislation affects in other ways the *economic* efficiency of the labor group.

Much labor legislation obviously has a direct relation to the personal health and efficiency of the wage-earners and their families. Safety and sanitary measures not only reduce the number of days lost during the year but also increase the general health and efficiency of the worker. Similarly, shorter hours will lessen fatigue and thus increase bodily resistance to disease. Moreover, shorter hours will make the workman more alert and energetic during the working hours. Twelve-hour-day workers cannot measure up in hourly output to eight-hour-day workers, and frequently the shorter workday shows an advantage in daily output. Social insurance helps to prevent physical deterioration of the worker and his family when wages cease because of sickness, accident, or unemployment. Minimum-wage laws increase the physical efficiency of the underpaid classes by enabling the worker to obtain the food, clothing, and other necessities of life requisite to a healthful existence.

So far as these measures serve to increase the personal efficiency of the wage-earning class, the marginal productivity of that class will be increased. So far as the productivity is increased, the benefits received from these measures are made economically possible and are paid out of the increased productivity. In the case of minimum-wage legislation the higher wages may thus come out of increased productivity. Social insurance need not result in any diminution of wages, provided productivity is increased sufficiently to pay for the cost of the insurance. If the employer is made to bear the expense of the insurance (as he is in workmen's compensation), it is clear that the wages paid in the long run cannot exceed the *net* marginal product of labor, that is, the amount remaining after the insurance expenses have been met. But the net marginal product need not be lower than before the insurance measure was put into effect, if there has been an increase in productivity. In the case of legislation on hours, the daily wage need not be reduced if the laborer produces as much in the shorter workday as he does in the longer day.

In the second place, the marginal productivity of labor may be increased by labor legislation even though there is no increase in personal

efficiency. The economic productivity of labor depends, it will be remembered, not only on the personal efficiency of the individual wage-earner but also on the relative supply of the different factors of production as well as the efficiency of the factors auxiliary to labor. The minimum-wage legislation and social-insurance measures may raise the marginal productivity of labor by exerting a powerful pressure on the employer to improve management and equipment to an extent which will make possible the payment of the higher wages and additional contributions required by law. As has been pointed out before, a high wage level is a powerful stimulant to improvements in management and technique. It is a challenge to the ingenuity of employers, and those who are not able to meet the test are likely to be squeezed out. There is a limit, of course, to such possible improvements. Nor is it possible for all firms to reach the maximum efficiency attained by superior firms. Hence legal minimum wages, and supplementary payments in the form of social-insurance contributions, if pushed very far, may easily reach a point too high for industry to sustain. In this event, if the law were really enforced (which would be difficult), a considerable number of the wage-earners to whom the law applied would be thrown out of work, since they would not be employed if they could not earn the wages and supplementary contributions required by law. The *active* labor supply would thus in effect be reduced, and hence the marginal product of labor would be raised to a point sufficiently high to pay the legal wages. In short, the total labor supply could not be employed at the abnormally high wage level.

Naturally the workers first thrown out would be the most inefficient workers. Thus the marginal product would be raised for two reasons: first, because the active labor supply would be reduced; second, because the elimination of the most inefficient would leave the level of personal efficiency higher than before.

A minimum-wage law and social-insurance measures may thus, for the time being, work a hardship on the least efficient workers, but in the long run even this class may gain. Among the inefficient workers who are thrown out of employment there may be many children and minors. A longer period for the education of the average child thus would be made possible, and if such education and training were fitted to his capacity, the efficiency of the next generation of workers could be increased. Moreover, all these measures serve to raise workingmen's concepts of what constitutes a proper standard of life. The effect of this psychological influence makes itself felt through parents' fore-

thought for the future of their children and probably thus indirectly affects the birth rate. As we have already seen, a restricted birth rate, within limits, is essential to the attainment and maintenance of a high standard of living. Furthermore, the throwing out of work of the most inefficient adult workers would probably compel society to give more liberal aid to defectives and partial defectives just below the margin of employability, with a view to educating and rehabilitating those who can be helped and caring properly for those who are hopelessly deficient. Moreover, legal minimum-wage laws may properly permit lower wage rates to persons who are partial defectives but nevertheless can do useful work and contribute to their own support.

PANACEAS AND THE LABOR PROBLEM

There are those who believe that the labor problem will finally be solved, not through trade unionism or labor legislation, but through revolution and the complete overthrow of the capitalist employer. Change is indeed the law of life, and though democratic countries are apparently moving in the direction of a sort of socialized capitalism, while in other parts of the world communism and fascism prevail, industrial classes continue to exist with open or submerged conflicts of economic interests. Private capitalism, modified by social control, persists tenaciously despite strains and frictions and periodic upheavals, which often serve to intensify the labor struggle. There is no final solution of the labor problem. It is a part of the eternal problem of human relations: of give and take between people engaged in a common enterprise. The labor problem is, in short, nothing more nor less than the art of living and making a living in co-operation with our fellow men. But this involves the adaptation of economic institutions to a changing world.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Suppose a bricklayers' union so strong that it would be able to raise wages to \$25 a day. Point out factors which would come into play making it unlikely that such high wages could be retained.
2. It is possible that a trade-union may raise wages without lowering profits or raising prices to the consumer. Under what conditions would this be possible?
3. A trade-union may raise wages, and the employers may pass the increased cost of production on to the consumers by charging higher prices. Under what conditions would this be likely to happen?

4. Under certain circumstances it is quite likely that a union by forcing the payment of higher wages would force marginal or typical employers out of business. What are these circumstances? What would be the effect upon consumers? Might such action result in social gain?

5. High wages of union men might result in lowering the wages of nonunion workmen. Under what conditions might this be the case? Would it necessarily be the case? Why not?

6. It is sometimes held that the rate of wages depends exclusively upon the relative bargaining strength of the two opposing parties. What do you think of this statement? Does the bargaining theory answer the question as to where wages come from and what sets the limits to the amount that can be paid in wages?

7. While a union may succeed in fixing wages at a very high level, it cannot force employers to hire the usual number of laborers at the exorbitantly high wage. What might the employers do to make it unnecessary to hire so many high-priced laborers? Might the buying public indirectly force the employer to hire fewer laborers? Explain.

8. Show how it might easily be possible for a certain class of wage-earners to suffer as a result of increasing their daily output. In view of this why does labor as a class generally gain from increased productivity?

9. The following table from W. I. King's *Earnings, Hours, and Employment in Prosperity and Depression* shows the decline in employment in the United States following the crisis of 1920. Note the industries that were hardest hit.

Maximum Cyclical Decline of Employment in 1920-1922 (Per Cent)

Agriculture	3.18
Extraction of minerals	29.66
Building and construction	18.92
Finance	7.14
Public and professional service	4.57
Domestic and personal service	4.11
Steam railways	29.68
Commerce and trade	2.78
All factories	29.97
Metals and metal products ¹	50.25

¹Vehicles, railroad cars, and all products not elsewhere recorded are included here.

CHAPTER XXIX · Some Dynamic Aspects of Distribution

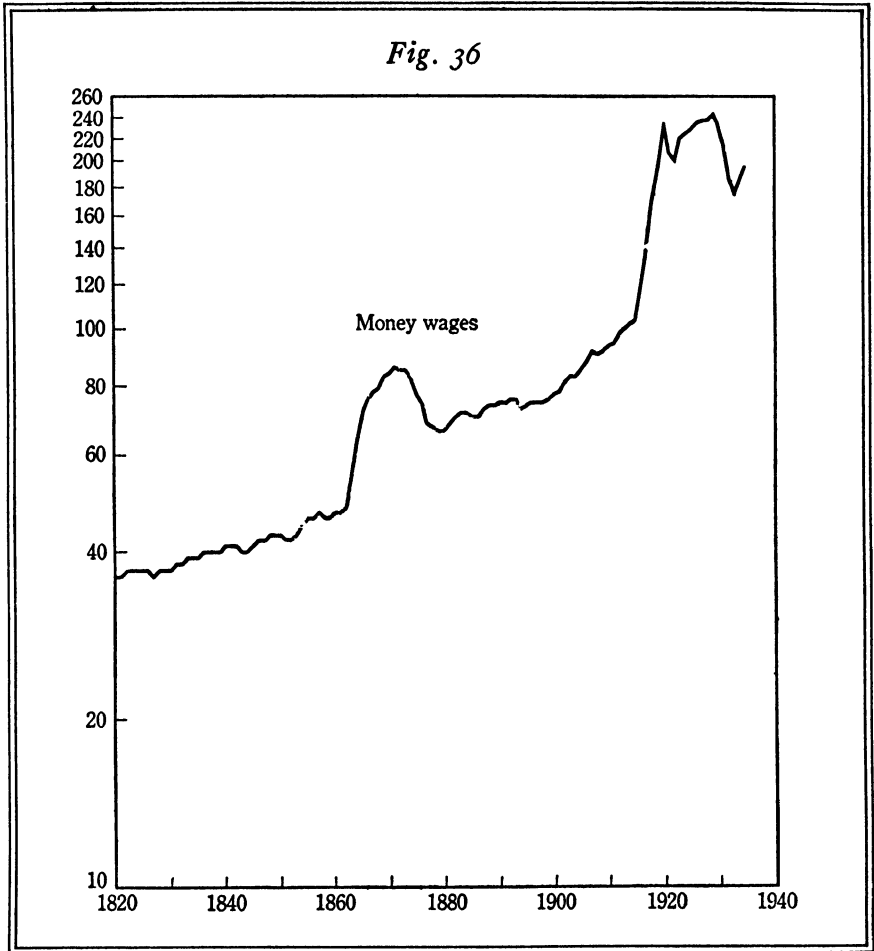


THE WAGE-EARNING CLASS; WAGES

The accompanying chart (Fig. 36) shows the historical trend of money wages¹ in the United States for over a century. Certain outstanding movements will be noted at once. Marked fluctuations appear at and after two great wars, the Civil War and the World War. These wage upheavals correspond to the great price upheavals, as will be seen by comparing the wage charts given here with the price chart in Fig. 27, page 322. It is obvious, therefore, that the general price level is one very important factor affecting the money-wage curve. If, however, we disregard these outstanding fluctuations we find that the general trend is steadily upward. No such tendency appears in the general wholesale-price curve. We conclude, therefore, that there are other very important factors operating upon the wage curve besides the movement of prices or the fluctuations in the value of money. If we reduce the wage curve to a purchasing-power basis, the effect of these other forces will become more clearly apparent.

In Fig. 37 is presented the movement of real, or purchasing-power, wages. The money-wage index before 1890 has been reduced to a purchasing-power basis by dividing it by an index of the wholesale prices of food, clothing, fuel, light, and house furnishings weighted roughly according to the expenditures of workingmen's families. This does not give an absolutely accurate index of real wages, even though the underlying data were perfectly accurate and representative. To get a more accurate index of real wages, of course, we should have to reduce money wages to a purchasing-power basis by means of a cost-of-living index. This would include rent as well as food, clothing, etc., and the prices should be at retail and not at wholesale. We have no adequate data on rent before 1913, and we have no continuous retail-price data, except for food, prior to 1890. Thus before 1890 wholesale

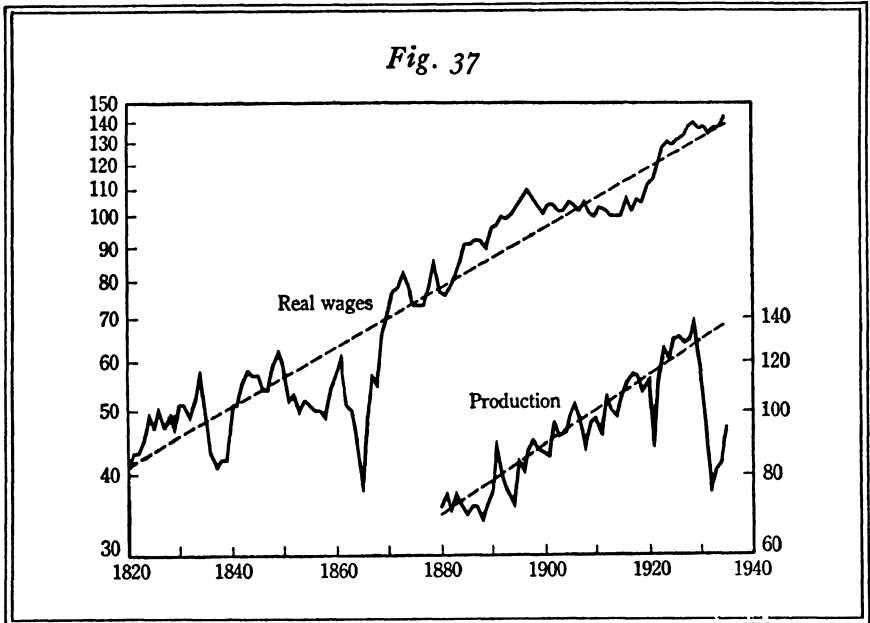
¹ Daily wages are used from 1820 to 1890; weekly wages, from 1890 to 1936. See Alvin H. Hansen, "Factors Affecting the Trend of Real Wages," *American Economic Review*, March, 1925. These data have been brought down to date.



prices have been used exclusively; from 1890 to 1913 the retail prices of food and the wholesale prices for other items have been used; after 1913 the cost-of-living index of the Bureau of Labor has been used. Wholesale prices deviated from retail prices only to a slight degree from 1890 to 1914. However, in periods of pronounced price upheaval wholesale prices fluctuate more violently than retail prices; hence when money wages are reduced to a purchasing-power basis by means of wholesale prices, the real-wage curve fluctuates more violently than is justified by the actual facts. It should moreover be noted that the true changes in real income cannot accurately be taken account of by means of even the most perfect price and wage data, since the character of goods consumed is constantly changing.

THE EFFECT OF PRODUCTIVITY ON WAGES

From the chart it is clear that apart from periodic movements, sometimes of longer and sometimes of shorter duration, the long-run trend of real wages has been decidedly upward. This would logically be expected in a century which witnessed a marvelous increase in production.



The chart shows a curve indicating the trend in the per capita production.¹ This is by no means the same thing as a curve showing the fluctuations in the marginal productivity of labor. The marginal product of wage-earners is not likely to increase as rapidly as the product per "gainful worker." As has been fully explained in a previous chapter, wages depend upon the marginal product of labor and not upon the average product. The leading cause of this increase in productivity per worker is the progressive improvement in the equipment with which each worker is supplied. The primary horsepower per worker in manufacturing increased about 90 per cent from 1899 to 1925. The

¹ Calculated in terms of the number of "gainful workers" in the country from decade to decade. "Gainful workers," as defined by the census, includes independent entrepreneurs, corporation officials, farmers, and professional people, as well as wage-earners. It includes all "workers" so defined, whether employed or unemployed.

See Proceedings of the Academy of Political Science (July, 1927), p. 7, and the National Industrial Conference Board Bulletins (December, 1927, p. 94, and Feb. 20, 1934, p. 10).

following table gives the increase in productivity in manufacturing from 1899 to 1933, together with more limited data in agriculture. Increased productivity may be expected to increase the real wages of the working class, especially if this increase affects products entering into the workingman's budget; however, if improvements result in lower cost and increased production of goods not purchased by workmen, the real incomes of other classes are increased, but there may be no gain for the wage-earners.

TABLE 27. Productivity per Worker

	MANUFACTURING		AGRICULTURE
	Output per Worker (1899 = 100)	Output per Man-Hour (1923-1925 = 100)	Output per Worker (1899 = 100)
1899	100		100
1909	110		106
1919	106		128
1923	134	95	
1924			147
1925	149	105	
1927	148	108	
1929	163	117	
1931	149	124	
1933	159	145	

Moreover, we have to consider the share of the national income which is distributed through public agencies. Much of the laborer's income does not appear in the pay envelope. An increasing proportion of the national income is being absorbed by the community as a whole, and distributed not according to the principles of value but according to the principles of need or social utility as interpreted by the play of political forces, whether in a democracy or in an authoritarian state. The wage-earner's pay envelope may or may not be affected by these public expenditures, depending on the manner in which the funds are raised. At any rate, these are matters that cannot be overlooked in a consideration of the relation between the movement of wages and national production.

THE EFFECT OF PRICE FLUCTUATIONS ON WAGES

The wage curves appearing in the charts are for the average *daily earnings* of employed workers before 1890, and for average *weekly earnings* after 1890. It should be noted that these curves do not show the wage

income for all workers as a class, employed and unemployed. The charts show only the wage income of *employed workers*. They do take account, however, of fluctuations in part-time and overtime work by employed wage-earners. In periods of prosperity, when prices tend to rise, there is less part-time employment and more overtime work, and in periods of depression the reverse is true. This factor tends to make the curves correlate directly with the short-run fluctuations of business and prices. However, the wage income of employed workers is affected not merely by fluctuations in part-time and overtime work, but also by changes in the wage rate per hour or per piece. Now these tend also to correlate directly with the prosperity cycle and with price movements. But it often happens that in periods when prices are rising, wage rates rise less rapidly than prices, and often in periods of price decline, wage rates lag behind. This is especially the case when the price movement is sharply upward or downward and for prolonged periods. The lag in wage rates behind pronounced price movements is frequently sufficient so that the *real* wage income of employed workers often correlates inversely with price movements. On the other hand, the total purchasing power of the entire wage-earning group regularly rises with prosperity and falls with depression, owing especially to increased employment in the former case and widespread unemployment in the latter.

Six major movements of prices and the earnings of employed workers are apparent. From 1820 to 1849 the tendency of real wages was upward and the trend of prices was downward. From 1849 to 1865 the general trend of prices was upward, and in this period real wages tended downward. From 1865 to 1897 the price trend was downward and the real-wage movement was strongly upward. From 1897 to 1919 the trend of prices was upward, but real wages remained, on the whole, stationary. With the stable prices of the nineteen-twenties real wages moved slowly upward. During the depression beginning in 1929, however, the earnings of employed workers fell more than the cost of living.

It is evident that price fluctuations have a powerful influence on the distribution of the national income. The marginal product of each factor determines the normal rate of return to each. But this normal distribution is constantly being upset, not only by changes in the relative supply and efficiency of the various factors but also by the redistribution of income accompanying changed prices. If prices rise, the benefit accrues immediately to the entrepreneurial class. In time, however, competitive bidding by entrepreneurs raises the prices paid for labor, land, and capital until the surplus profits are absorbed. In the mean-

time a fresh rise in prices has created a new margin of profits above costs. Entrepreneurs naturally do not pass these gains on to the other factors until compelled to do so by the pressure of competition. It is quite usual, therefore, in rising-price periods, for wages and other costs to lag behind prices. Rising prices amount, in fact, to a redistribution of the national income in favor of the entrepreneurial class. They amount to an enforced taxation of wage-earners, salaried persons, investors, and landlords with long-term rent contracts.

From 1897 to 1913, and again from 1922 to 1929, when real wages remained substantially stationary in spite of a great increase in national production, business profits increased enormously. This shows itself, for example, in the rise in the stock market. From 1897 to 1913 railroad and industrial stocks advanced nearly 100 per cent, while from 1922 to 1929 common-stock prices advanced two or three fold.

Moreover, in the period from 1897 to 1920, large corporate surpluses were being set aside out of profits, and this development in part resulted from the rising-price trend. A large part of the increased national production was in the field of producers' goods.¹ Important net additions were made to our capital equipment. This new capital came immediately out of the savings of the corporations, but in reality in large measure out of the enforced savings imposed by the rise in the price level upon wage-earners, salaried people, bond and mortgage holders, insurance-policy holders, etc. Rising prices cause a gap between the marginal productivity of the various factors employed by the entrepreneur and the return that each receives. Indeed, in such periods it is literally true that labor (as well as certain other factors) does not receive the full value of its product.

It may well be, indeed, that certain of the groups enumerated above enjoy a larger income in spite of the enforced saving. Thus though many wage-earners find their real-wage rates (hourly or piece rates) reduced in rising-price periods, yet their real annual earnings are larger because of the increased employment. But that does not obviate the fact that they are none the less taxed by the rise in prices, as is shown by the reduction of their real-wage rates. They get less pay for the work done, but they get larger real incomes by doing more work. Other groups are less fortunate. Salaried people, bond and mortgage holders, and insurance-policy holders suffer a net loss because of the decline in the purchasing power of their fixed money incomes.

¹ E. E. Day and W. M. Persons, *Review of Economic Statistics*, November, 1920, and July, 1927.

During these periods of easy profits the large capital investment of the corporations increased the equipment with which the laborers worked. This added capital was constantly raising the marginal productivity of labor. The added capital contributed to a higher standard of living in another way. It has been historically true that periods of intense prosperity have also been periods of rapid technical progress. Corporations that have ample incomes are much more likely to set aside large sums for research than those that can barely meet fixed charges and pay small dividends. Improved techniques result, in the long run, in a rise in the standard of comfort for all classes, including a majority of the working classes. The opportunity for large profit stimulates speculation and wildcat ventures, it is true, but it also stimulates discovery and the foundation of many new enterprises. It must not be forgotten that the period from 1897 to 1920 was the period that saw the development of the automobile, and the further improvement and spread of the use of electricity and electrical appliances, all of which have contributed to the pleasure and well-being of practically all classes in America.

On the other hand, labor may get more than its marginal product in falling-price periods. Prices may fall, but labor clings tenaciously to its achieved wage rates and does not hesitate even to suffer unemployment in preference to wage reductions. Therefore in such periods the employer is hard put to it to reduce costs in some other way than by reducing wage rates. This he may do by eliminating waste, by introducing improved processes and machinery, and by better management.

Moreover, the marginal product of labor is higher in falling-price periods because of the reduction in the labor force. Thus the high real-wage rates are purchased, at least in part, at the price of an excessive volume of unemployment.

The low profits of falling-price periods result in a diminution in corporate surpluses and social saving. The production of consumers' goods assumes a relatively greater importance. Thus, in the falling-price period industry operates to a large extent on the capital accumulated in the previous period of rising prices.

Farmers suffer particularly in periods of falling prices. During the last century and a quarter we have had, along with many movements of less importance, five outstanding periods of severe decline in the purchasing power of farm products, beginning with the following dates: 1817, 1837, 1857, 1919, and 1929. These were all periods in which the general price level was falling heavily. What is the reason for this

relationship? When general prices fall, industrial depression prevails, and the more precipitate the fall the more severe is the depression. When industry is depressed, the market for farm products is poor for two reasons: first, in such periods, manufacturing consumes a smaller quantity of the raw material produced on farms (about three fourths of the raw material used in manufacture comes from farms); and secondly, the purchasing power of wage-earners, much of which is applied to farm products directly or indirectly, is low, because of the prevailing unemployment. In spite of this restricted market, farmers continue to produce approximately the normal supply of the raw materials of industry, while the current supply of other goods and services is rigidly controlled. In a period of depression wage-earners do not engage to sell all their labor at whatever price it will fetch; they prefer unemployment to a precipitate decline in wages. Entrepreneurs curtail production in the face of a falling market. This curtailment checks the extent of the decline in the selling price of manufactured goods. The farmers, however, largely unable to control supply, find themselves in the position of residual claimants. Throwing their commodities on the market in undiminished volume while other groups are controlling supply, they face an adverse purchasing-power ratio. It would be interesting to know how low wage rates would have fallen in 1932 if wage-earners had insisted on selling the entire supply of labor at whatever rates it might fetch.

During the great depression that began in 1929 there was an extraordinarily high correlation between the gross income of farmers and the total money wages received by urban wage-earners. Both fell by 1932 to about one third of the 1929 level, and both recovered at substantially similar rates beginning with 1933. The fluctuation in the gross income of farmers was almost wholly due to the decline and subsequent rise in the prices of farm products. The fluctuation in urban pay rolls was due mainly to the decline and subsequent improvement in employment but in part also to changes in wage rates. The close correlation of gross farm income and urban pay rolls is evidence of the interrelation between industrial prosperity and the prices obtained by farmers.

It is sometimes argued that high wage rates in the depression period are favorable to farmers, for it is thought that the high rates will create a good market for farm products. But unfortunately high wage rates do not help an unemployed man to buy commodities. Widespread unemployment causes low urban purchasing power in the depression period even though wage rates are relatively high, and in consequence the farmer suffers.

THE EFFECT OF IMMIGRATION ON WAGES

The failure of real wages to keep pace with productivity in the two decades preceding 1914 can be explained in part by the unprecedented wave of immigration which began at the turn of the century and continued until the outbreak of the war. In six of the ten years from 1905 to 1914 over a million immigrants arrived annually. The labor market was constantly fed with a new supply of applicants for jobs, making the competitive situation favorable to employers and unfavorable to wage-earners. Under these circumstances it is not surprising that real wages lagged behind the increase in productivity.

The restriction of immigration in more recent years has been a potent factor working toward an increase in real wages. European immigration was almost completely stopped by the war, and in the postwar period restrictive legislation has kept the volume at relatively low figures. In 1921 an act was passed, chiefly applicable to Europe, which allowed each country an annual quota equal to 3 per cent of the number of its nationals in this country in 1910, thereby limiting the immigration from Europe to about 358,000. This act continued in force until 1924, when a new law was passed, reducing the quotas to 2 per cent of the foreign-born in this country in 1890. Accordingly the maximum total European immigration was reduced to about 165,000. This measure especially reduced the number from southern and eastern Europe, since relatively few natives of these countries lived in the United States in 1890. Under the act of 1924 the quota for southern and eastern Europe was only about 20,000.

The 1924 act provided that after three years the per centum method of restriction should be superseded by the "national origins" method. This clause in the act, however, was postponed from time to time until it finally went into effect on July 1, 1930. According to this method 150,000 were to be admitted each year, and of this total each foreign nation was assigned a share corresponding to the percentage of the total population of the United States in the year 1920 which could be attributed, in terms of "national origin," to that country. Thus it was estimated that about 44 per cent of the population of the United States in 1920 were of British origin, and so Great Britain was allowed 44 per cent of the total of 150,000, or about 66,000 immigrants per year. Germany was allowed nearly 26,000, the Irish Free State nearly 18,000, and Scandinavia nearly 7000. The quota restriction did not apply to any of the countries of North and South America.

During the great depression, immigration from European countries fell far below their quotas. The vast unemployment in the United States tended to prevent immigrants from coming, and, moreover, the American immigration authorities prevented the immigration of any person who was likely, in a period of industrial depression, to become a public charge.

The decline in immigration during the nineteen-twenties was particularly marked in the case of unskilled workers. This followed naturally from the drastic quota limitations that were imposed upon the countries of southern and eastern Europe from which this class of immigrants largely came. For northern Europe the quotas were more generous, and the skilled workers from these countries could still migrate to the United States in considerable numbers. During the depression, however, even this migration virtually ceased. Accordingly, a shortage of skilled workers in certain lines had already made itself felt in 1936 with the revival in industrial activity. American industry must develop a more adequate apprenticeship system for the training of skilled workers.

Table 28 indicates the broad changes in immigration since 1910. The vast net immigration from southern and eastern Europe during prewar years had been converted, even in predepression years to a minus quantity, the outflow to these countries having exceeded the small number admitted under the quotas. On the other hand, the net immigration from northern and western Europe had been reduced to only two thirds of the prewar level. The Canadian immigration was three or four times greater than in the prewar period, and Mexican immigration two or three times the former level. During the depression years, taking all countries together, there was a small net outflow.

TABLE 28. Average Annual Net Immigration, in Thousands

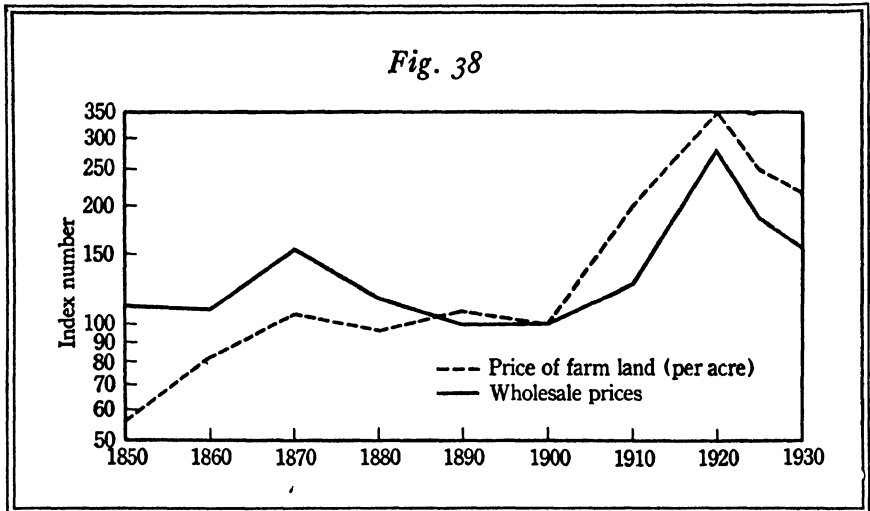
PERIOD	MEXICO	CANADA	NORTHERN AND WESTERN EUROPE	SOUTHERN AND EASTERN EUROPE
1910-1914	17	27	157	513
1915-1919	10	57	26	- 3
1920-1924	47	101	110	95
1925-1927	34	89	107	- 13
1928-1930	31	66	} 110	
1931-1934	- 17	9		

THE EFFECT OF UNIONISM ON WAGES

The cost of living fell heavily after 1920, but money wages after a moderate decline returned approximately to the peak of the war period. Unions unquestionably were a factor of more or less importance in preventing drastic reduction in wages during the deflation of 1920-1922 and again in 1929-1932. The resistance of the trade-unions was particularly evident in certain industries, notably textiles, transportation, mining, clothing, and printing. Moreover, between 1933 and 1937, aided partly by the NRA and partly by the revival of business, wage rates were restored to the 1929 level.

THE FARMING CLASS; LAND VALUE AND LAND RENT

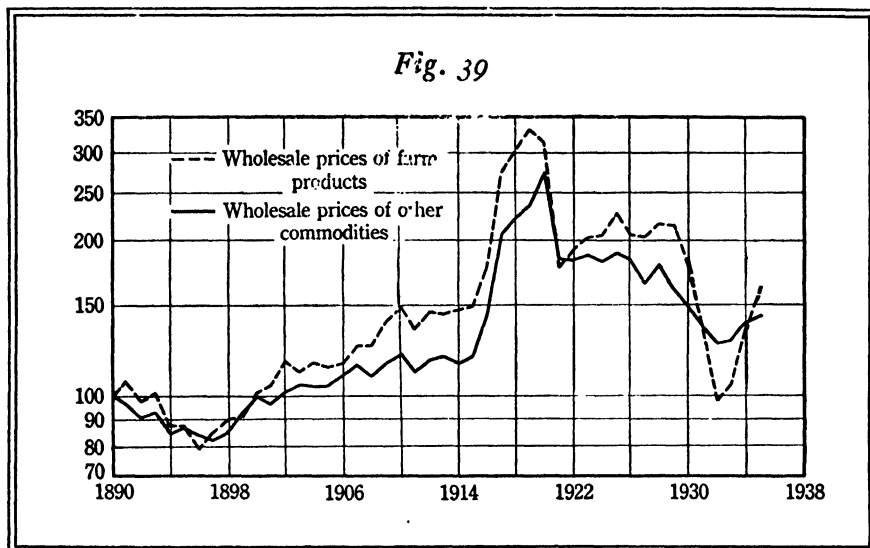
We turn now to a consideration of the fluctuations in the value of farm land. Fig. 38 gives a curve showing the average price of farm land per acre in the United States (improvements included). The value of land



is determined by the capitalization of its rental value. The capitalization will vary, not only with the rental value, but also with the rate of interest. The curve here given is defective, since by taking the average price per acre of land there is a progressive change in the thing evaluated, new land having been added continually and some old land abandoned. Nevertheless the curve does probably indicate broad general tendencies.

From 1850 to 1870 the trend of land values was strongly upward, from 1870 to 1900 it was slightly downward, from 1900 to 1920 there was a phenomenal rise in land values, and after 1920 there was a sharp decline.

The curves (for decades, and for five-year periods after 1920) in Fig. 38 indicate clearly that a considerable part of the movement of land values is due to the fluctuations in the value of money. Land, like commodities generally, tends to rise as money depreciates in value and to fall with general prices as money appreciates in value.



But this by no means tells the whole story. From 1900 to 1910, land values doubled while general prices increased only about 25 per cent. It is clear that the general increase in prices, or, what amounts to the same thing, the depreciation of money, is not sufficient to explain this phenomenal rise in value. To understand this remarkable increase, we must distinguish between the movement of prices of farm products and the prices of other commodities. These movements are shown in Fig. 39. It will be noted that the prices of farm products increased 47 per cent from 1900 to 1910, that the prices of other commodities (not including farm products) increased only 18 per cent, and that the general price index (including farm products) increased only 25 per cent. The disappearance of free land, technological developments, and the continued growth of population were factors in this change. The prices of farm products increased more than two and a half times as much in this

decade as the prices of other commodities and nearly twice as much as the general wholesale-price index, which included agricultural and non-agricultural products.

The effect of these movements on land rents will become clear from a simplified illustration. This illustration does not take account of all the factors present, but it does isolate some of the more important elements in the situation. Let us suppose that a farmer in 1900 sold the products of his farm for \$2000 and that his expenses (not including interest on the investment in the land) were \$1000. The rental value of his land was then \$1000. By 1910 the same crop would sell for \$2940. Let us suppose that his expenses increased in proportion to the increase in the general price level.¹ They would then increase from \$1000 to \$1250. The net rental value of his land is now \$1690. The rental value of his land thus has increased 69 per cent whereas the prices of farm products increased only 47 per cent. It is thus evident that in a period in which the prices of farm products are rising more rapidly than expenses, the rental value of land will increase even more rapidly than the prices of farm products.

But while the rental value of land increased, according to our illustration, 69 per cent from 1900 to 1910,² land values doubled. How is this divergence between the rental value and the capitalized value of land to be accounted for? The answer is to be found in the fact that the purchasers of farm land anticipated a continued increase in the rental value of land. The prospective buyer capitalized not only the present rental value of the land but also the anticipated future increase in the rental value. The prospective buyer of land therefore was willing to pay a price for the land greater than was warranted by the prevailing rental value. And it must be noted that a relatively small annual increase in the rental value accumulates very rapidly. Thus, if the present rental value of a piece of land is \$1000 and the anticipated annual increase in the rental value is \$50, in twenty years the rental value will be doubled. If we assume a constant rate of interest,³ a \$50 annual increase in the rental value anticipated indefinitely will make the land, now yielding \$1000 rent, worth very much more than if there were no

¹The price of farm machinery increased about 15 per cent in this period and the wages of farm labor increased about 40 per cent. Many other factors affected farm expenses, some favorably, such as the introduction of improved machinery, fertilizing grasses, and improved seeds.

²The increase in cash rents was considerably less than the increase in the economic rent assumed above. Compare C. R. Chambers, *Relation of Land Income to Land Value*.

³In point of fact the interest rate rose during the last six years of the decade 1900 to 1910.

expectation of any increase at all. How much more, will depend upon the rate of interest at which the increased rental value is capitalized.¹

From 1900 to 1910 the capitalized value of land increased more than the rental value because of the anticipated future increases in rental value; and the rental value increased more than the prices of farm products because of the fact that expenses did not keep pace with the selling price of the products. Here, then, we find the set of interrelated forces which produced the enormous increase in the value of land from 1900 to 1910. The increase would have been still greater had not the favorable factors been offset in part by rising interest rates which tended to reduce the capitalized value.

From 1910 to 1920, land values for the country as a whole did not keep pace with the general rise in prices, though in some states they did. This means that the full increase in rental value was not capitalized, since it was not expected that the great rise in agricultural commodity prices incident to the war period could be counted on to be permanent. Nevertheless people relied on the permanence of this increase to a far greater extent than was justified, as they learned to their sorrow when the drop in farm-product prices came in the latter half of 1920.

The fall in wholesale prices of farm products was far greater in 1921 than the price decline of other commodities. The prices received by the farmers *at the farm* fell still more in proportion. These facts, together with the continuance of high taxes, caused a severe shrinkage of the rental values of farm land, in many cases reducing them to absolutely nothing. There followed a heavy reduction in land values from 1920 to 1930.

It is clear from this discussion that one of the important factors affecting the value of land is the movement of the prices of farm products in comparison with the general price level; in other words, the trend of the purchasing power of farm products. The movements of the purchasing power of farm products during the last century will be of interest in this connection.²

Table 29 shows the purchasing power of farm products (that is, the prices of farm products in terms of general prices) by decades from 1790

¹The formula by which this may be calculated is as follows. Let \$1000 equal the present rental value, \$50 the anticipated increase in rental value, and 4 per cent the rate of interest. Then

$$\frac{1000}{.04} + \frac{50}{(.04)^2} = 25,000 + 31,250 = \$56,250.$$

²See Alvin H. Hansen, "The Effect of Price Fluctuations on Agriculture," *Journal of Political Economy*, April, 1925.

to 1935. Except for the decades 1820-1829, 1860-1869, and the period 1920-1935, the movement has been steadily upward. This rise in the purchasing power of farm products is remarkable and indeed somewhat surprising in view of the fact that the nineteenth century witnessed a phenomenal development of agriculture, both extensive and intensive. In that century the railroads opened up enormous tracts of virgin land. But manufacturing developed even more rapidly. The factory system and large-scale production matched the flood of farm products with a still greater outpouring of manufactured goods. So far as the United States is concerned, agriculture expanded enormously up to about 1890, but when we consider the world market it is clear that manufacturing production outran agricultural production in relation to the changes in demand for these two classes of products. After 1890 agricultural production slackened, whereas manufacturing developed apace; consequently the purchasing power of farm products rose.¹

TABLE 29

DECADE	PURCHASING POWER OF FARM PRODUCTS, WEIGHTED INDEX (1913 = 100)	DECADE	PURCHASING POWER OF FARM PRODUCTS, WEIGHTED INDEX (1913 = 100)
1790-1799 . . .	48	1870-1879 . . .	79
1800-1809 . . .	51	1880-1889 . . .	84
1810-1819 . . .	55	1890-1899 . . .	88
1820-1829 . . .	51	1900-1909 . . .	93
1830-1839 . . .	62	1910-1919 . . .	104
1840-1849 . . .	63	1920-1929 . . .	97
1850-1859 . . .	82	1930-1935 . . .	89
1860-1869 . . .	78		

For an illustration of these movements, compare the prices of cloth and clothing during the last century with the prices of farm products. As Table 30 shows, clothing was very expensive in the early part of the century and farm products were relatively cheap. As the century progressed, clothing fell in price in relation to farm products, because of the technical progress in the textile and clothing industry.

Relative to the prewar (1913) relationship, the prices of farm products in the greater part of the postwar period have ruled low compared with industrial prices. This is due to a combination of circumstances.

¹ Account must be taken not only of changes in the demand for agricultural and industrial products, but also of the effect of changes in cost of production on supply.

During the war, when food supplies were scarce in Europe, acreage and production were greatly expanded in the United States, Canada, Australia, Argentina, and other areas. After the war, agricultural production was restored to its former level in the war-ravished countries of Europe and indeed raised still further by a system of high protection. At the same time technological developments in the use of farm machinery further extended production in the new countries. In conjunction with these developments on the supply side, the demand situation was far from favorable. The rate of population growth was rapidly slowing down, the per capita demand for cereals was declining, owing partly to the introduction of more varied diets and partly to the decline (with less heavy manual labor) in the number of calories consumed. Moreover the gasoline engine had supplanted some 7,000,000 horses and mules which formerly had consumed large quantities of feeds grown on the farm. All these factors combined to produce huge agricultural surpluses, which began after 1925 to overhang the market. While these surpluses had been materially reduced by 1936, in part owing to deliberate restrictions of output in various countries, and especially to several serious droughts, it remains true that the world's agricultural plant is overdeveloped. This situation presents a serious problem for agriculture throughout the world and tends to keep agricultural prices low in relation to the prices of industrial products.

TABLE 30

DECADE	WEIGHTED INDEX OF THE PRICE OF FARM PRODUCTS	WEIGHTED INDEX OF THE PRICE OF CLOTH AND CLOTHING
1820-1829	53	171
1830-1839	67	161
1840-1849	59	124
1850-1859	77	122
1860-1869	107	197
1870-1879	87	135
1880-1889	74	104
1890-1899	65	83
1900-1909	81	91
1913	100	100

But the fluctuation in the purchasing power of farm products is not the only factor affecting the farmer's real income. The farmer is a property-owner and, as such, the fluctuations in the value of money may make or break him. During the decade 1923-1933, for example, it

was not the decline in the *purchasing power* of farmers' incomes that was mainly responsible for the foreclosure sales and bankruptcies; it was the high price paid for land at the peak of the war boom. Or, to put it in another way, the seat of the trouble was the fact that the farmers' money income (not purchasing-power income) had not maintained its 1916-1920 level, upon which expectancy the land was bought and mortgaged. Even had the purchasing-power income remained as great as in those years, the farmers' difficulty nevertheless in many cases would have been very great. Overvaluation of land in view of the subsequent money incomes was the great evil.

From 1880 to 1895 the trend of the purchasing power of farm products was slightly upward, but the money value of farm products was falling greatly. The real income of the farmers was not falling in this period. Yet it was a period of "hard times" for farmers. As property-owners they were being subjected to a constant process of confiscation, accomplished through the constant fall in prices.

The ambition of practically every American gainfully employed in agriculture is to become a landowner, and farmers who are already landowners naturally put their savings into the only kind of investment they know anything about, namely, land. Rarely is the purchaser of land able to advance the entire purchase price; mortgage indebtedness is therefore the lot of a considerable proportion of farm-owners. Let us suppose that a farmer bought a piece of land for \$20,000 in 1867 and that he mortgaged it for \$10,000. By 1878 the prices of farm products had fallen to just half the 1867 level. Then there was a considerable rise in the early eighties, followed by a decline to still lower levels in the late eighties and nineties. The trouble was not with the *purchasing power* of the farmer's money income but with the *number of dollars* in his income. The number of dollars had severely declined. The fewer dollars would buy just as much in current goods, but unfortunately the mortgage did not shrink along with other prices. The farmer received fewer dollars for his products and paid fewer dollars for his clothes, horses, machinery, lumber, etc., but he still had to pay the same number of dollars for his mortgage. If the farmer had not been in debt his real income would have been as favorable as in any previous period. Certainly the purchasing power of his products averaged as high as ever before, with the single exception of the prosperous decade of the fifties. But being compelled to pay interest on the mortgage, which refused to shrink along with other money values, his real net income was in fact constantly being reduced, with the decline in

the size of his money income. Moreover, the decline in prices brought with it, as has been shown, a decline in the price of land per acre. The farm-owner's equity was being squeezed out. A 50 per cent fall in the price of land would confiscate all the property of a farmer who had paid cash on half the purchase price of his land.

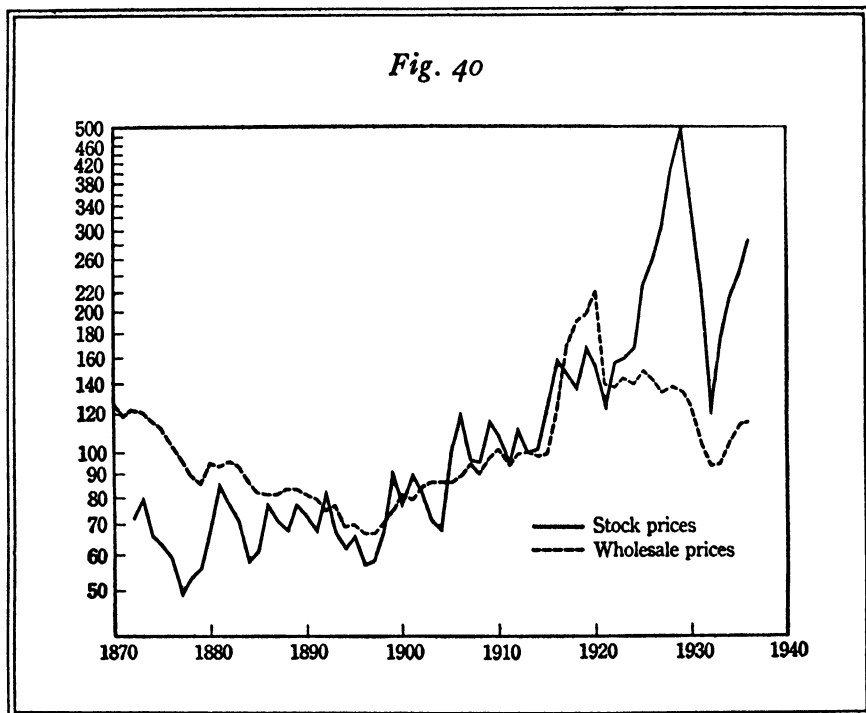
Nor is a period of rising prices altogether an unmixed blessing to the farming population. True, in such periods the mortgages dwindle away relative to the rising money incomes, but the evil of overvaluation appears. In such periods land is valued not on the basis of its present income but on the expectation that its rental value will continue to rise. In the period before the World War land was frequently purchased for twice the value warranted by the current rental value. Thus while prospective purchasers, renters who were struggling to climb up into the landowning class, might hope that their money incomes would increase with the rise in prices, the advantage of this was already discounted in the excessive price paid for the land at the outset. It was therefore mainly the rising price level from 1897 to 1920 that caused the overvaluation of land under which farmers subsequently suffered. When the anticipated increase in farm-product prices failed or, worse yet, was converted into a precipitate decline such as we witnessed in 1920-1922 and again in 1929-1932 the farmer was left stranded high and dry with his highly capitalized land. The marked rise in prices since 1933 has in considerable measure relieved this discrepancy.

Even though the general price level were stabilized, the prices of farm products might still fluctuate. However, if extreme inflation and deflation of general prices could be eliminated there would not arise such violent discrepancies between rental value and capitalized value as have harassed farmers over the past century.

THE BUSINESS CLASS; PROFITS

The statistical data with reference to business profits are considerable for the last few years, but do not extend back very far in time. Excepting those for the banks and the railroads, there are no data of any consequence extending over any considerable period of time. We have the index numbers of the prices of stocks back to 1872. Fig. 40 shows the relative prices of industrial common stocks from 1872 to 1936, together with the movement of wholesale prices.¹

¹The figures for 1872-1896 are from the *Commercial and Financial Chronicle*; for 1897-1936, from Dow-Jones and Standard Statistics.



The wholesale-price trend falls more rapidly from 1872 to 1896 than the stock-market trend. Probably money profits declined more nearly in proportion to the general price level; but interest rates also were falling, and hence the capitalized value of these earnings would decline less rapidly than the money earnings. The upward trend of industrial stock prices from 1896 to 1914 corresponds closely to the upward movement of wholesale prices for the same period. But there are certain divergences that should be noted. It is especially noteworthy that industrial stocks advanced 85 per cent from the low point in 1896 to the period 1905-1913, whereas wholesale prices advanced only 40 per cent. This was clearly a period in which huge profits were being made. Business profits were outrunning the general price level in this period, in part because costs (interest charges, wages, etc.) were lagging behind selling prices. The advantage of these lagging costs accrued to the entrepreneurial class. Moreover, this was a period of rapid technical progress, and the first effect of such improvements is to enhance pure profits. Furthermore, new industries, such as the electrical industry and the public utilities, were developing. New industries which prove successful are likely to yield high rates of profit.

TABLE 31. Dividends and Interest Payments of Corporations in the United States, in Billions of Dollars

YEAR	TOTAL DIVIDEND PAYMENTS (CASH)	INTEREST PAYMENTS
1922	3.4	3.1
1923	4.2	3.3
1924	4.3	3.4
1925	5.2	3.6
1926	5.9	4.0
1927	6.4	4.4
1928	7.1	4.6
1929	8.4	4.9
1930	8.2	4.9
1931	6.2	4.5
1932	5.9	4.0
1933	3.1	3.5

TABLE 32. Net Incomes of Corporations in the United States, in Billions of Dollars

YEAR	TOTAL NET INCOMES OF ALL CORPORATIONS REPORTING NET PROFITS	TOTAL DEFICITS OF ALL CORPORATIONS REPORTING NET LOSSES
1916	8.8	0.7
1917	10.7	0.6
1918	8.4	0.7
1919	9.4	1.0
1920	7.9	2.0
1921	4.3	3.9
1922	7.0	2.2
1923	8.3	2.0
1924	7.6	2.2
1925	9.6	2.0
1926	9.7	2.2
1927	9.0	2.5
1928	10.6	2.4
1929	11.7	3.0
1930	6.4	4.9
1931	3.7	7.0
1932	2.2	7.8
1933	3.0	5.5
1934	4.3	4.1

At the time of the World War the stock market advanced rapidly in response to higher prices and profits. But the period of greatest advance came in the late twenties. Large profits were made despite the high wages paid. The indication is that the productivity per em-

ployee greatly increased. This increase was accomplished largely through the introduction of more and better machinery and through better management. Thus from 1923 to 1929 there was no correlation between the movements of stock prices and those of wholesale prices. During these years there was a phenomenal advance in stock prices, whereas wholesale prices remained practically stationary. This advance in the stock market is to be explained partly by the rapid introduction of cost-reducing improvements, noted above, which resulted in a great increase in business profits; partly from the great development of new industries, including the automobile, electrical-equipment, chemical, and motion-picture industries; and partly by the unusually large volume of funds available for investment and speculation. These funds came in part from the large profits which were distributed to stockholders or retained in the business and in part from the abundant supply of bank credit which, with the huge gold reserves, the banking system was easily able to provide. The abundance of funds available for investment and speculation insured a low rate of interest.¹ With rising profits and a low rate of interest, the prices of stocks inevitably rose rapidly. A part of the great increase in the value of industrial shares in the late twenties was thus the result of high corporate earnings capitalized at low rates of interest. A part was due to a wave of speculative activity based on the expectation of further advances in security prices.

TABLE 33. Average Net Income of 10,020 Corporations for the Three Years 1911-1913²

RATIO OF NET INCOME TO INVESTED CAPITAL (PER CENT)	NUMBER OF CORPORATIONS		NET INCOME	
	Actual	Per Cent of Total	Actual (in Thousands)	Per Cent of Total
Under 6	1,368	13.7	\$19,011	3.4
6 to 8	737	7.4	26,770	4.8
8 to 12	2,415	24.1	168,246	30.5
12 to 20	2,929	29.3	163,091	29.5
20 to 30	1,366	13.6	93,004	16.8
30 to 40	553	5.5	35,430	6.4
40 to 50	233	2.3	17,817	3.2
50 to 75	259	2.6	16,656	3.0
75 to 100	75	0.7	7,606	1.4
Over 100	85	0.8	4,719	0.9
	10,020	100.0	\$552,350	99.9

¹ With the rapid increase in incomes we should expect a reduction in the marginal cost of saving.

² David Friday, *Profits, Wages, and Prices*, p. 41.

Table 33 shows the prewar earnings of 10,020 corporations. It will be noted that over 20 per cent earned less than 8 per cent on invested capital (these were marginal firms making no pure profits whatever), that 25 per cent earned from 8 to 12 per cent, and that 30 per cent earned from 12 to 20 per cent. The majority of these corporations therefore earned from 8 to 20 per cent. If the 12 to 20 per cent group were divided into two groups, one ranging from 12 to 16 per cent and the other from 16 to 20 per cent, making them strictly comparable to the 8 to 12 group, it would probably appear that the model group is the 8 to 12 group. A little over 25 per cent earned more than 20 per cent.

TABLE 34. Business Failures in the United States, by Decades

DECADE	PERCENTAGE OF BUSINESS CONCERNS FAILING ANNUALLY	TOTAL LIABILITIES OF BUSINESS FAILURES ANNUALLY (IN MILLIONS)
1880-1889 . . .	0.96	\$132
1890-1899 . . .	1.10	178
1900-1909 . . .	0.90	146
1910-1919 . . .	0.88	218
1920-1929 . . .	0.99	498
1930-1935 . . .	1.06	555

TABLE 35. Percentage of Corporations in the United States Reporting Net Profits and Deficits

YEAR	PERCENTAGE REPORTING NET PROFITS	PERCENTAGE REPORTING NET DEFICITS	PERCENTAGE INACTIVE
1916	60.7	39.3	
1917	66.0	34.0	
1918	63.6	36.4	
1919	65.5	34.5	
1920	58.8	41.2	
1921	48.0	52.0	
1922	55.5	44.5	
1923	58.5	41.5	
1924	56.6	43.4	
1925	58.7	41.3	
1926	56.7	43.3	
1927	54.7	34.9	10.4
1928	54.2	35.3	10.5
1929	52.9	36.6	10.5
1930	42.7	46.6	10.7
1931	34.1	55.0	11.0
1932	16.2	72.6	11.2
1933	21.8	66.9	11.3

But not all corporations earn a net profit. An index of business losses may be found in the record of business failures. Table 34 shows the percentage of business firms failing and the total liabilities of business failures from 1880 to 1933. Table 35 gives the percentage of corporations reporting net profits and also of those reporting deficits from 1916 to 1934. Of the 236,000 corporations reporting a net income in 1924, 76 per cent earned less than \$10,000, 20 per cent between \$10,000 and \$100,000, and only 4 per cent over \$100,000. This latter group, however, earned over 75 per cent of the total net income reported by all these corporations.

THE INVESTING CLASS; INTEREST RATES

We next turn to the movement of interest rates. Interest rates on long-time investments are best shown by the yield of high-grade bonds. The bond yields on ten high-grade railroad bonds, reduced to index numbers, are shown in Fig. 41.¹

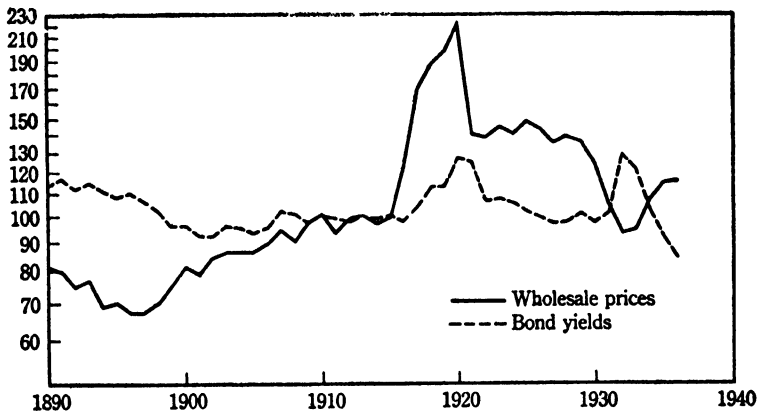
It will be noted that three long-run trends of interest rates from 1890 to 1936 are evident from this chart. From 1890 to 1901 the trend was downward, from 1901 to 1920 it was upward, and from 1920 to 1936 it was downward. To some extent the trends correspond to the movement of the wholesale-price level. Prices, however, reached bottom in 1897, whereas interest rates reached bottom in 1901, four years later. Moreover, the interest rate increased only about 41 per cent from 1901 to 1920 while wholesale prices increased 186 per cent, or over four times as much. Finally, attention should be called to the fact that by 1927 the interest rate had fallen nearly to the prewar level, whereas prices had been substantially stabilized at 50 per cent above the prewar level. This illustrates the fact that the rate of interest is not affected by the *level* of prices. While prices were *rising*, interest rates also rose; with the fall in prices in 1921, interest rates also fell. But when prices became stabilized after 1922, interest rates continued to fall until they approached the prewar normal. This accords with the theory of interest explained in Chapter XXVI. Interest rates are temporarily affected by monetary factors, but in the long run they are controlled by time preference and by the productivity of capital.

The peculiar banking situation in the United States since 1914 accounts in part for the great lag of interest rates behind prices from 1914

¹ By courtesy of the Harvard Economic Society; see *Review of Economic Statistics*, 1919 and 1923; for 1913-1936, Standard Statistics, 60 domestic bonds.

to 1920 and also for the rapid decline of interest rates from 1920 to 1929. The establishment of the Federal Reserve System increased banking efficiency, making possible a larger expansion of bank credit on a given reserve of gold. Moreover, during the war period the large gold imports greatly increased the volume of gold reserves, and this tended to retard the advance of interest rates. Similarly, in the postwar period the enormous gold supply has produced a favorable credit situation which has contributed to the fall in the rate of interest.

Fig. 41



During the great depression, bond yields fluctuated considerably. In the period of sharp liquidation of assets resulting from foreign withdrawals of funds, from bank failures, and from business bankruptcies, bond values declined sharply and therefore yields rose. When the banking system was reorganized in 1933 and the process of liquidation ceased, the accumulation of investment funds, together with the huge surplus bank reserves incident to gold imports, drove the yield down to the very low figure of less than 3 per cent on government and high-grade corporation bonds.

We have noted that while the rate of interest is wholly independent of the *level* of prices, nevertheless with a *rise* in prices interest rates also tend to rise; likewise interest rates tend to fall with a fall in prices. This general relationship is not difficult to explain. A falling-price

period is favorable to fixed investors. The stipulated return on the bond or mortgage remains fixed as to the *number* of dollars received at the same time that each dollar received is increasing in value. During such periods a stockholder is likely to receive a lower money income, since the earnings of business concerns are likely to fall in declining-price periods. Landowners likewise are likely to receive declining money rents from their land. Investments therefore tend to be transferred from stocks and land to the bond market. Hence bonds are bid up in price, and as a result the yield declines.

Conversely, with rising prices, a bondholder is the loser. On the other hand, the earnings on stocks, as we have seen, increase greatly in such periods, as do also the rents from land. Many bondholders readily perceive that it is more advantageous to be a stockholder or a landowner in periods of rising prices. The tendency is to sell the bonds and transfer the investment to stocks or real estate. Bond prices therefore fall in value and, conversely, the bond yields rise.

It is easy, however, to exaggerate the gain or loss enjoyed or suffered by the investor class because of price fluctuations. Let us take an extreme illustration. Suppose a man lends \$1000 for a year at 3 per cent. A year from now he will get back \$1030. If no change occurs in the price level in the meantime, the lender will receive back 3 per cent more purchasing power than he lent. This constitutes his interest. Suppose, however, that prices double during the course of the year. It now requires \$2000 to equal the purchasing power of \$1000 before the rise. To receive back a purchasing power over goods and commodities equivalent to the \$1030, on the basis of the old price level, the lender will now have to get \$2060. In other words, the interest rate would have to be 106 per cent in order to yield a 3 per cent return on the real purchasing power lent. However, the investor group as a class are not likely to apply the principal to the purchase of goods and commodities; more probably they reinvest it in other loans — bonds and mortgages. Now while commodity prices rise interest rates also rise, and so bond prices fall. Hence in periods of rising prices the purchasing power of the investor's principal is not falling but rising.

Let us suppose that the interest rate doubles along with prices. The interest rate will then rise until, at the close of the year during which prices doubled, the interest rate will be 6 per cent. As a result of this, perpetual bonds (that is, bonds that have no date of maturity) paying a stipulated 3 per cent rate will fall to approximately half their previous price, say from \$1000 to \$500. Now let us see what will be the situation

with our investor who has made a \$1000 loan for one year at 6 per cent on the assumption that the price level would rise 100 per cent during the year. He will get back \$1060. The \$60 is his income from the investment and the \$1000 is his principal. Our investor has obviously suffered no loss in real income, since his \$60 interest will now (prices having doubled) buy as much as his former \$30 annual interest would purchase. Nor has he suffered any loss as a property-holder. He can take his \$1000 and buy two \$1000 bonds (par value), each paying 3 per cent interest. The market value of these two bonds, to be sure, is only about \$500 each,¹ because 3 per cent bonds with no maturity date will be worth only about half as much as 6 per cent bonds currently floated.

It may appear from this argument that the investor class would neither gain nor lose if the interest rates received by them did indeed rise and fall in proportion to the rise and fall of prices. But, as a matter of fact, such is not the case. The owner of a bond or mortgage receives a fixed money return even though the current interest rate may be changing. If the investor holds a bond or mortgage with a distant maturity, there is no possible way in which he can increase his money income even though interest rates are rising as rapidly as prices. So long as he holds his bond or mortgage he receives, of course, only the former stipulated rate. If, on the other hand, he sells it in order to invest his money in a bond that yields a higher current rate, he finds that he loses on the principal in the same proportion as he gains on the stipulated rate of interest received. As a result, his money income is no higher. If, for example, he is getting 3 per cent and the current rate is 4 per cent, his bond, formerly worth \$100 when capitalized at 3 per cent, is now worth only \$75 when capitalized at 4 per cent. Seventy-five dollars invested at 4 per cent will yield no higher return than \$100 invested at 3 per cent.

Only in the event that the bond or mortgage matures can the investor take advantage of the higher current interest rate. If he holds it until maturity he will get back his principal, which may be reinvested at the higher current rate.

We may, then, lay down this general principle: investors, since their income is fixed during the life of the loan, suffer in rising-price periods a loss in real income directly proportional to the fall in the purchasing power of their incomes. If, however, the current rate of interest has risen proportionately with prices, the investor is able to recover his former purchasing-power status every time his loan matures. In point of fact,

¹This illustration probably exaggerates the decline of perpetual bonds, since it is likely that a future decline in the interest rate will be forecast.

of course, it is not necessary for him to wait until maturity to recover a part of his purchasing-power status, since the value of a maturing bond depends partly on the date of maturity.

This situation, however, would obtain only so long as prices and interest rates continued to rise in equal proportions. Should the price level become stabilized at the higher level, interest rates would again fall to the normal level established by the demand for and the supply of capital. But the investor whose bond matures at this juncture finds that he is compelled to reinvest his principal at the lower rate of interest in spite of the fact that prices remain at their former high level. His real income is thus seriously reduced. Moreover, even while prices are in process of rising, the rate of interest does not rise *proportionally* with prices. Thus bondholders usually lose heavily in rising-price periods and make large gains in falling-price periods.

OTHER FACTORS AFFECTING THE RATE OF INTEREST

There are, however, other factors that affect the fluctuations in the rate of interest, besides the depreciation of money and the fluctuation in prices. The high rates of contract interest that obtained before 1880 can be accounted for partly by the greater risk that industry faced in a new country with an unstable, fluctuating, and speculative market. Moreover, the country was undeveloped and the demand for additional capital, particularly for buildings and railroads, was very great. The combined influence of the fall in the price level, the stabilization of the market as population and business became more settled, and the decline in the great drafts made upon capital by railroad and building construction, led to the decline in the rate of interest to the low point reached at the close of the nineteenth century. Then the great development of public utilities, the electrical industry, the automobile industry, road-building, modern houses and office buildings, and finally the war, greatly increased the demand for capital, and this, together with the upward movement of prices, led to the great rise in the rate of interest witnessed during the first quarter of the twentieth century.

Probably building construction, including houses, factories, and other buildings, absorbed over one half of our savings for the decade of the nineteen-twenties; over one tenth went into additional automobiles and trucks; about one fifth went into new railroad equipment and such public utilities as electric light and power, telegraph, and telephone; about one fifteenth went into highway construction and pavements; and

about one twentieth into foreign investments. Altogether we were saving in that decade about one sixth of our national income, or from twelve to fifteen billion dollars a year.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The average weekly earnings of employees in New York state factories were as follows :

	MARCH, 1928	NOVEMBER, 1936
Male shop employees	\$34.29	\$28.83
Female shop employees	19.39	15.90

2. The following were approximately the hourly earnings received by various grades of labor in the United States in November, 1936.¹

Skilled and semiskilled (factory male labor)	\$0.70
Unskilled (factory male labor)	0.50
Female (factory labor)	0.43
Common (construction labor)	0.58
Skilled (construction labor)	1.18
Steel industry (wages per hour)	0.48

3. Wages paid in manufacturing constituted the following percentages of the value of all manufactured products in the years given:

1909	16.6
1914	16.8
1919	16.8
1921	18.8
1923	18.2
1925	17.1
1931	17.4

What bearing do these facts have on the degree to which wage changes affect profits? Do they throw any light on the relative efficiency of labor in different periods?

4. The total wage payments made in New York state factories reached the index number of 264 in 1920 compared with 100 for the prewar period. Average weekly earnings, however, rose only to 222. In 1925 both stood at the figure 223. How can these movements be explained? Why were total wage payments so much higher in 1920 than average weekly earnings?

5. The real wages of clerical and lower-salaried workers are approximately equal to their prewar level, whereas the real wages of manual workers are from 25 to 30 per cent above the prewar base. How do you explain these differences?

6. "The secular increase in population would seem to offer every prospect in the long run for an increase in economic rent and in the price of land. If and

¹Survey of Current Business.

when this occurs, then the real wages of city workers must be lowered unless such a tendency be counterbalanced by an equivalent or more than equivalent increase in the productivity of manufactures.”—DOUGLAS. Argue for or against the proposition that the interests of urban wage-earners and farmers are opposite.

7. “From 1920 to 1924 there was a heavy decline in the membership of trade unions. In this very period there was a rise in real wages. It is therefore evident that trade unions cannot claim any credit for the gain made.” Do you agree? Why, or why not?

8. The National Transportation Institute gives the following figures on the relative growth in number of persons employed, in capital investment, and in physical output from 1900 to 1920. The figures are for the year 1920, and are based on 1900 as 100.

INDUSTRY	PERSONS ENGAGED	CAPITAL	OUTPUT
Agriculture	104	381	138
Manufacturing	208	495	228
Mines	170	302	231
Railroads	199	193	292

Rank the industries according to (1) increase in capital invested per employee; (2) gain in output per employee. Do you see in these facts any application of the principles of distribution?

CHAPTER XXX · Inequalities of Property-Ownership and Income

We have seen in the preceding chapter that the economic condition of the wage-earning class has greatly improved during the last century. But though the masses are rising in the economic scale, it might well be that the rich are gaining still more. Is property-ownership more highly concentrated in the hands of the very rich now than formerly? Is the middle class being wiped out? Do the poorer classes receive as large a proportion of the total national income now as formerly? How much of the national income goes to the very rich? If these huge incomes were distributed to other classes, how much would they gain? It is not possible to answer these questions with entire satisfaction. We are, however, sufficiently sure of some of the facts to indicate the broad general tendencies.

Inequalities in the ownership of property are in some measure connected with inequalities in income, but not so closely as at first thought might be supposed. As a matter of fact, the distribution of income is much more nearly equal than the distribution of property. There are many professional and salaried persons who earn large incomes but who own little or no property. Moreover, a large part of the national income is paid to wage-earners who own relatively little property. On the other hand, farmers, for example, own considerable property but receive relatively small incomes.

The distribution of income is important from the standpoint of consumption, but the distribution of consumption does not necessarily correlate closely with the distribution of income. In point of fact, many wealthy persons spend only a small percentage of their incomes on personal consumption. Who, then, gets the benefit from these incomes? The wealthy persons in question may spend their surplus incomes on objects for public consumption, such as libraries; they may establish philanthropic foundations; or they may invest in capital goods. Persons with large incomes therefore are in a position to *direct* the currents of public consumption and to *control* production. Inequalities in consumption, of course, follow from unequal distribution of income; but the concentration of control that goes with large incomes and fortunes is

probably of even greater significance, for it raises grave questions with respect to democratic institutions and social control. Have these large fortunes and surplus incomes directed consumption and production more wisely than would have been possible with greater democratic control? Does a man like Henry Ford control the consumption of the nation to the extent that he controls its productive processes? Is the democratic control of income and consumption more important than the democratic control of property and production? Is it likely that a democracy can more wisely and successfully socialize its income than its property? Can it achieve genuine control of the one without the other, or are they connected in such a manner that control of the one involves control of the other?

It is not the purpose of this chapter to attempt to give a complete answer to these questions. Rather, this chapter will give the student a statistical analysis with respect to the distribution of income and wealth, in the belief that such data will at least help to illuminate these problems.

DISTRIBUTION OF WEALTH IN MASSACHUSETTS

Table 36 shows the distribution of estates among men over twenty-five years of age who died in Massachusetts in the three periods 1829-1831, 1859-1861, and 1889-1891. This material has been compiled from the Twenty-fifth Annual Report of the Massachusetts Bureau of Labor Statistics.

The number of men over twenty-five years of age who die in any one year greatly exceeds the number of estates probated in any one year. The reason for this divergence is that considerably over half the adult males who die have little or no property and hence no estate to be probated.

The distribution of wealth among deceased males, of course, is not an absolutely accurate index of the distribution of wealth among the living. It is well known that a considerable amount of wealth escapes the probate courts. Many wealthy men transfer much of their property while they are still alive to members of their family or to charitable, educational, and other similar institutions. All such gifts, transferred during the lifetime of the donor and therefore not included in his will, escape the probate courts. Moreover, life-insurance policies which are assigned to a designated beneficiary, and not to the estate, escape the probate records. Finally, a considerable amount of property is held by joint title of husband and wife, and such property passes to the surviving party without any action in probate court. Attention should be called

TABLE 36. Distribution of the Probated Estates of Males in
 Massachusetts¹

PERIOD	ESTIMATED ² PROPORTION OF ADULT MALES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
1829-1831	52	Under \$1,000 \$1,000 to \$5,000 \$5,000 to \$25,000 \$25,000 to \$100,000 \$100,000 and over	47.5 36.6 13.6 1.9 0.4	3.5 20.2 30.4 22.7 23.2	\$4,352
1859-1861	55	Under \$1,000 \$1,000 to \$5,000 \$5,000 to \$25,000 \$25,000 to \$100,000 \$100,000 and over	31.7 40.9 21.6 4.3 1.5	1.5 11.3 24.5 22.9 39.9	\$8,985
1889-1891	60	Under \$1,000 \$1,000 to \$5,000 \$5,000 to \$25,000 \$25,000 to \$100,000 \$100,000 and over	24.6 40.9 25.9 6.2 2.3	0.8 7.4 20.6 22.1 49.1	\$13,658

also to the fact that the age distribution of living adults, of course, is very different from the age distribution of those who die in any given year. The proportion of young men would be much greater among those living. As a rule, men accumulate more property as they grow older. The proportion of men owning property of over \$100,000 would probably be less among those living than among those who die in any given year. The average age of deceased males is greater than the average age of those living, and it is likely that the inequality in the ownership of wealth is greater in a given population with an average age of sixty years than in one with an average age of only thirty years.

In 1890 about 60 per cent of the men over twenty-five years of age who died in Massachusetts left no estate to be probated. Of those whose estates were probated, about 25 per cent left less than \$1000 worth of property, about 40 per cent left between \$1000 and \$5000, about 25 per cent left between \$5000 and \$25,000, and less than 10 per cent left over \$25,000.

¹ See the Twenty-fifth Annual Report of the Massachusetts Bureau of Labor Statistics.

² Estimated from United States census data on population and mortality.

These figures are likely to give an erroneous impression to the present generation, because the purchasing power of a dollar is considerably less today than it was in 1890. A thousand dollars' worth of property today means much less than a thousand dollars' worth of property in 1890. The three periods given in the table, however, are substantially comparable, since the purchasing power of money was not widely different in these periods.

Certain broad conclusions can be drawn from the figures presented in the table. In the first place, it should be noted that in the sixty-year interval from 1830 to 1890 the proportion of men who died leaving no estate to be probated increased from 52 per cent in 1830 to 60 per cent in 1890. Thus there was no indication that the great propertyless masses were becoming property-owners. Rather the tendency appeared to be slightly in the opposite direction.

On the other hand, if we center our attention on that portion of the male population who were owners of property in the three respective periods, 1830, 1860, and 1890, we find that the average property-owner in 1860 owned over twice as much as the average property-owner in 1830 and that the average property-owner in 1890 owned over three times as much as the average property-owner in 1830; in other words, the property-owners in 1890 were much wealthier than the property-owners one or two generations before.

Now how was the wealth distributed among the different *property-owners*, rich and poor? We are excluding entirely from our consideration here the majority who owned no property whatever except a few personal belongings and who therefore had no estates to be probated; we are centering our attention upon the property-owners, those leaving estates to be probated. What changes took place in the relative property holdings of these men? Do the data show any tendency toward greater concentration of wealth in the hands of a few, or is there a tendency toward greater equality of holdings among all property-owners?

In the first place, it is clear from Table 36 that property-owners were progressively moving up into higher valuation groups, leaving relatively fewer in the lowest group (those owning property worth less than \$1000). In 1830 nearly half the property-owners left an estate valued at less than \$1000. In 1890 less than one fourth left estates of this value. The rest had moved up into the higher-valuation groups. The \$1000-\$5000 group increased slightly in proportion, but the upper groups gained the most. The \$5000-\$25,000 group was proportionally nearly twice as large as in 1830, the \$25,000-\$100,000 group was over three times as large, and the

Inequalities of Property-Ownership and Income 543

“over \$100,000” group was over five times as large. This does not mean, however, that property was passing increasingly into the hands of a few rich. What had actually taken place was that property-owners all along the line had become wealthier.

This fact is brought out in Table 37. In this table the property-holders (deceased males with probated estates) have been divided more or less arbitrarily into five groups: the richest group constituting 1 per cent of those leaving estates, the rich constituting 4 per cent, the upper middle group 15 per cent, the lower middle group 30 per cent, and the poorest property-owners 50 per cent of the total. It will be noted from the table that the richest 1 per cent owned 35 per cent of the property, both in 1890 and in 1830. There is no evidence, therefore, of increasing concentration in the hands of a few. At the other extreme, the poorest 50 per cent owned substantially the same proportion of the total wealth in 1890 as in 1830. On the other hand, the moderately rich gained relatively at the expense of the middle groups. The rich 4 per cent owned 26.5 per cent in 1890 as against 22 per cent in 1830, whereas the middle groups, constituting 45 per cent of the total, owned 34 per cent in 1890 as against 39 per cent in 1830.

TABLE 37. Distribution of Total Value of the Probated Estates of Males in Massachusetts

PROPERTY-HOLDING GROUP, MALES	PERCENTAGE OF TOTAL PROPERTY PROBATED WHICH WAS HELD BY EACH GROUP		
	1829-1831	1859-1861	1889-1891
Richest group, constituting 1 per cent of estates probated	35	34.5	35
Moderately rich, constituting 4 per cent of estates probated	22	26.5	26.5
Upper middle group, constituting 15 per cent of estates probated	25	21.0	23
Lower middle group, constituting 30 per cent of estates probated	14	13.5	11
Poorest group, constituting 50 per cent of estates probated	4	4.5	4.5

The reader should bear clearly in mind the difference between the *relative* and the *absolute* position of these property-owners. From the absolute standpoint the average value of the property left by members of each of the classifications given was roughly three times as great in 1890 as in 1830; but the gain made by the middle groups was slightly

TABLE 38. Distribution of Probated Estates of Males and Females

PLACE AND PERIOD	ESTIMATED PROPORTION OF ADULTS OF BOTH SEXES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
Massachusetts 1829-1831	71	Under \$1,000	51.2	4.2	\$3,919
		\$1,000 to \$5,000	34.5	20.9	
		\$5,000 to \$25,000	12.2	30.6	
		\$25,000 to \$100,000	1.8	22.8	
		\$100,000 and over	0.3	21.6	
Massachusetts 1859-1861	68	Under \$1,000	35.3	2.0	\$7,694
		\$1,000 to \$5,000	40.8	25.0	
		\$5,000 to \$25,000	18.8	23.0	
		\$25,000 to \$100,000	3.8	37.3	
		\$100,000 and over	1.2	12.8	
Massachusetts 1889-1891	65	Under \$1,000	27.1	1.2	\$10,649
		\$1,000 to \$5,000	42.4	9.7	
		\$5,000 to \$25,000	23.7	23.7	
		\$25,000 to \$100,000	5.1	22.6	
		\$100,000 and over	1.7	42.9	
United States 1912-1923 ¹	57	Under \$1,000	24.9	0.8	\$15,428
		\$1,000 to \$5,000	37.6	6.4	
		\$5,000 to \$25,000	27.5	19.4	
		\$25,000 to \$100,000	7.7	22.9	
		\$100,000 and over	2.2	50.4	

less than that made on the average by all property-owners (the rich gaining slightly more than the average), and the gain made by the richest and poorest groups was almost exactly equal to that made on the average by all property-owners. The general conclusion is that the distribution of wealth among property-owners was substantially the same in 1890 as in 1830. On the other hand, the *absolute amount* of property held by each group was about three times as great in the later period.

These conclusions are somewhat modified if we include in our data the estates of both males and females. The data are given in Tables 38 and 39. It will be noted that the proportion of adults of both sexes who died leaving no estate to be probated *decreased* from 71 per cent in

¹W. I. King, *Journal of the American Statistical Association* (June, 1927), p. 145.

TABLE 39. Distribution of Total Value of the Probated Estates of Males and Females in Massachusetts and the United States

PROPERTY-HOLDING GROUPS, MALES AND FEMALES	PERCENTAGE OF TOTAL PROPERTY WHICH WAS HELD BY EACH GROUP			
	Massachusetts (1829-1831)	Massachusetts (1859-1861)	Massachusetts (1889-1891)	United States (1912-1923)
Richest group, constituting 1 per cent of estates probated	34.5	34.5	35	37
Moderately rich, constituting 4 per cent of estates probated . . .	22.5	25.5	25.5	25
Upper middle group, constituting 15 per cent of estates probated . . .	24	23	22.5	22
Lower middle group, constituting 30 per cent of estates probated . . .	15	12.5	12.5	11.5
Poorest group, constituting 50 per cent of estates probated . . .	4	4.5	4.5	4.5

1830 to 65 per cent in 1890 in Massachusetts, and 57 per cent in 1912-1923 in the United States as represented by twelve widely scattered states and the District of Columbia. It will be remembered, on the other hand, that the proportion of adult males who died leaving no estate *increased* from 52 per cent in 1830 to 60 per cent in 1890. Thus while the proportion of propertyless males was increasing slightly, the proportion of propertyless adults of both sexes was decreasing. Are we to conclude that there were therefore fewer propertyless *families* in 1890 than in 1830? It is difficult to say. It is clear that there was a markedly growing tendency for women to own property. But if this merely means, for the most part, that in 1890 the estate passed to the widow, was held by her for a few years, and was then passed on to the next generation, whereas in 1830 the estate passed directly to the sons, such a situation would not indicate any increase in family well-being. If, on the other hand, an increasing proportion of women, both married and unmarried, held property independently of their husbands or other male relatives, it is clear that this would imply a net gain. Probably both assumptions are true in part, and hence the net conclusion would seem to be that the propertyless class did not change proportionally to any considerable extent either up or down during the period in question.

Other tendencies are substantially the same, whether we take both sexes or males alone. The increase in the average value of estates does not appear so great if both sexes are considered. The upward-climbing tendency of property-owners generally is apparent in substantially the same manner as already discussed with reference to males. Similarly, the proportion of the wealth held by the rich groups, the middle groups, and the poorest group is not appreciably modified by the inclusion of females. The tendencies are the same as have already been indicated.

DISTRIBUTION OF WEALTH IN TWO MINNESOTA COUNTIES

Table 40 gives the property distribution of deceased persons for the year 1923 in Freeborn and Meeker counties, both in Minnesota.¹ These data are fragmentary and, of course, cannot be compared closely with the Massachusetts data. They are, however, interesting illustrations of the inequality in the distribution of wealth in farming districts in the Middle West. These counties are essentially rural, the largest city being Albert Lea, with a population of 11,000. By constructing a Lorenz curve from the data given in the table, we find that the richest 1 per cent owned 8 per cent of the wealth, the moderately rich 4 per cent owned 17 per cent of the wealth, the upper middle group (15 per cent) 34 per cent, the lower middle group (30 per cent) 29 per cent, and the poorest group of property-owners (50 per cent) 12 per cent of the property probated. Moreover, more than half left no property at all as far as the probate records show. From the figures just given it will

TABLE 40. Distribution of the Probated Estates of Males and Females in Freeborn and Meeker Counties, Minnesota

PERIOD	PROPORTION OF ADULT MALES AND FEMALES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
1923	51	Under \$1,000	10.0	0.5	\$10,809
		\$1,000 to \$5,000	33.3	7.8	
		\$5,000 to \$25,000	47.3	52.1	
		\$25,000 to \$100,000	8.7	33.4	
		Over \$100,000	0.7	6.2	

¹ This material was made available by the courtesy of Judge J. W. Wright of Litchfield, Minnesota, and Judge A. W. Johnson of Albert Lea, Minnesota.

be noted that the middle groups owned about two thirds of the property probated. This large percentage is to be expected in a distinctly rural community with no large fortunes.

Dr. Willford I. King found that the distribution of wealth among the living is considerably less unequal than among deceased persons. According to his estimate the richest 1 per cent of the gainfully employed population own about one third of the wealth, and the richest 1 per cent of adult deceased persons (including those with no probated estate) owned over half the wealth of those who have died. The richest 2 per cent of the gainfully employed own about 40 per cent of the wealth, but the richest 2 per cent of adult deceased persons owned 62 per cent of the wealth possessed by those who have died. Ten per cent of the gainfully employed own 64 per cent of the wealth, and 10 per cent of the adult deceased persons owned 86 per cent of the total wealth owned by these individuals.

DISTRIBUTION OF INCOME

Table 41 gives the total income for the people of the United States by decades from 1850 to 1910 and for certain years from 1910 to 1936, together with the per capita income for these years.

It is in the division of the income, however, that we are particularly interested. Table 42 gives the distribution of personal incomes as shown by the Federal income-tax returns for the years 1918, 1924, and 1930. It should be noted that the number of incomes in the \$1000-\$2000 class is

TABLE 41¹

YEAR	NATIONAL INCOME (IN BILLIONS OF DOLLARS)	PER CAPITA INCOME
1850	2.2	\$95
1860	3.6	116
1870	6.7	174
1880	7.4	147
1890	12.1	192
1900	18.0	236
1910	30.5	332
1922	65.6	597
1929	80.8	665
1933	41.8	332
1936	63.8	497

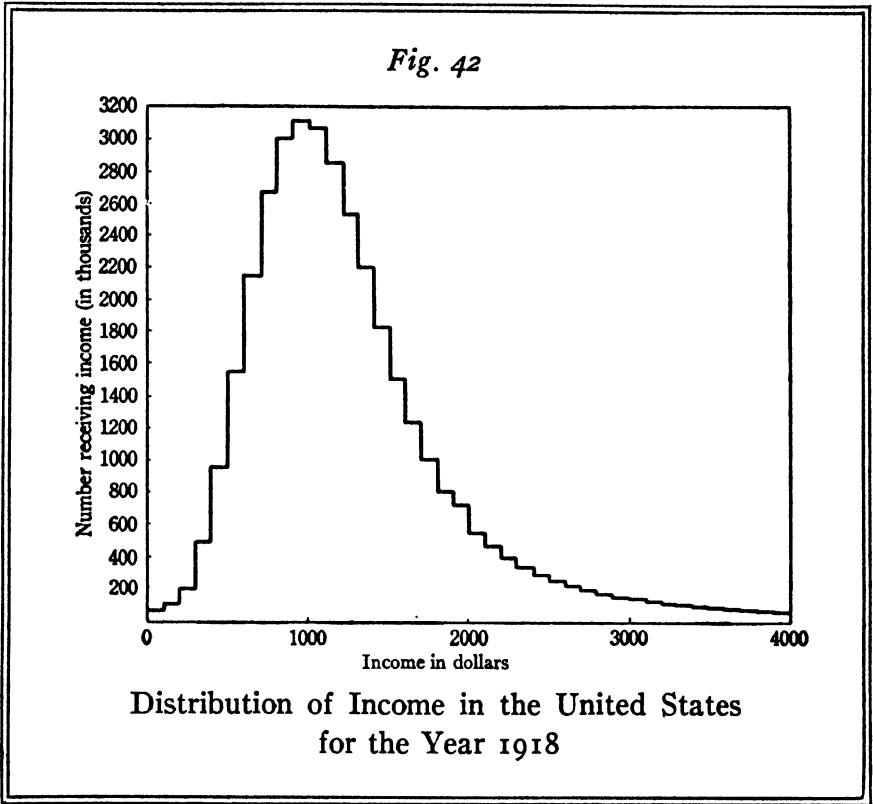
¹ 1850-1910 from W. I. King, *Wealth and Income of the United States*; 1922 from News-Bulletin of the National Bureau of Economic Research, Feb. 21, 1927; 1929-1936 from Survey of Current Business, June, 1937.

much greater than is indicated here, since the income-tax returns for this class include only unmarried persons or married persons not living together. Moreover, the number of returns in the three periods are not comparable, since the exemptions have been changed from time to time.

TABLE 42. Income Distribution as Shown by the Federal Income-Tax Returns

YEAR	INCOME CLASS	NUMBER OF RETURNS	NET INCOME (IN MILLIONS OF DOLLARS)	AVERAGE TAX RATE (PER CENT)
1918	\$1,000 to \$2,000	1,516,938	2,232	1.19
	\$2,000 to \$3,000	1,495,878	3,627	0.98
	\$3,000 to \$5,000	932,336	3,535	2.35
	\$5,000 to \$10,000	319,356	2,146	4.34
	\$10,000 to \$25,000	116,569	1,737	8.20
	\$25,000 to \$50,000	28,542	978	13.32
	\$50,000 to \$100,000	9,996	680	21.69
	\$100,000 to \$150,000	2,358	284	33.68
	\$150,000 to \$300,000	1,514	305	44.64
	\$300,000 to \$500,000	382	145	54.77
	\$500,000 to \$1,000,000	178	119	58.65
\$1,000,000 and over	67	137	64.65	
1924	\$1,000 to \$2,000	2,413,881	3,564	0.29
	\$2,000 to \$3,000	2,112,993	5,277	0.19
	\$3,000 to \$5,000	1,800,900	6,828	0.39
	\$5,000 to \$10,000	437,330	2,991	0.96
	\$10,000 to \$25,000	191,216	2,855	2.73
	\$25,000 to \$50,000	47,061	1,600	6.84
	\$50,000 to \$100,000	15,816	1,067	12.81
	\$100,000 to \$150,000	3,065	378	20.04
	\$150,000 to \$300,000	1,876	375	24.69
	\$300,000 to \$500,000	456	171	26.73
	\$500,000 to \$1,000,000	242	158	26.87
\$1,000,000 and over	75	156	30.27	
1930	\$1,000 to \$2,000	909,155	1,495	.08
	\$2,000 to \$3,000	767,684	1,864	.18
	\$3,000 to \$5,000	1,070,239	4,152	.13
	\$5,000 to \$10,000	550,977	3,724	.47
	\$10,000 to \$25,000	198,762	2,923	1.70
	\$25,000 to \$50,000	40,845	1,384	5.25
	\$50,000 to \$100,000	13,645	919	9.51
	\$100,000 to \$150,000	3,111	374	13.03
	\$150,000 to \$300,000	2,071	419	14.91
	\$300,000 to \$500,000	552	207	15.96
	\$500,000 to \$1,000,000	318	212	15.20
\$1,000,000 and over	150	360	16.98	

Table 42 includes, however, only about 10 per cent of all the income-receivers. To get a more complete picture of the distribution of the income of all the gainfully employed, it becomes necessary to make estimates based on a wide variety of sources. These include interest on tax-exempt bonds, tax-exempt salaries, the imputed rent of homeowners, farm products consumed directly by farmers, annual earnings



of the various classes of wage-earners, income of wage-earners derived from sources other than wages, and the incomes of the lower-salaried classes. From a wide variety of records the National Bureau of Economic Research has pieced together the available data and constructed an income-distribution table for the year 1918. The above chart (Fig. 42) shows this distribution of income among the various income-receivers in the United States for that year. It includes everyone receiving an income, men, women, and children (with the exception of soldiers and sailors and marines). To some extent it exaggerates the

number of persons receiving low incomes, since farmers' incomes are larger than the money value here assigned to them. Thus farm products used by the farmer's family are valued at the price the farmer could get for them rather than at what he would have to pay for them. Moreover, a farm yields a considerable number of perquisites the value of which it is impossible to measure accurately. Again, some considerable part of the real income of the average wage-earner's family is produced directly and never enters into the valuation of the market, for example the work of the housewife, plumbing, and other repairs about the house performed by the members of the family.

The chart gives a picture of the money incomes of income-receivers, but it does not give an accurate picture of the real distribution of the national income among individuals, for the reason that part of the taxes paid out of certain incomes to the government are redistributed back again by the government to individuals (often not the same as the taxpayers) in various kinds of services. The income of every family is increased, for example, by the services of free schools, free parks, free libraries, museums, paved streets, etc. Certain important taxes, such as the Federal income and inheritance taxes, special assessments, and taxes on property used for business purposes are not deducted from the incomes here given. It follows, therefore, that the net incomes of the well-to-do, who pay these taxes, are smaller than is indicated in the chart and the incomes of the poorer classes are somewhat larger than is indicated, since a part of their income comes to them in the form of free governmental services and assistance.

From the chart it will be noticed that the modal income in 1918 was about \$1000. By modal income we mean the income received by the largest number of persons. At first thought it may seem to some that this is contrary to common observation and knowledge. Yet it must be remembered that we are here including all income-receivers, whether men, women, or children. In New York State factories the average weekly earnings in 1918 were \$20.35. This would amount to \$1058 for fifty-two full weeks. But we all know that few wage-earners work fifty-two full weeks in any year. These figures are cited merely to show the inherent reasonableness of the finding of the National Bureau of Economic Research that the modal income in 1918 was \$1000.

It will be seen from the chart that the number of persons receiving the given income rises rapidly up to \$1000 and then diminishes until at \$4000 the income curve is rapidly approaching the base line. The chart includes only incomes ranging from zero to \$4000. If our chart,

drawn on the scale here given, were to include all incomes, it would have to be extended about a mile and a half to include Henry Ford's income of about \$120,000,000 a year. Only about three fourths (78 per cent) of the national income is included in the chart; the remaining fourth (22 per cent) would be included in the narrow but exceedingly long area which we may imagine if the chart were drawn large enough to include the total national income.

Thus 22 per cent of the total national income is included in the imaginary portion of our chart, and this share is received by but 3 per cent of the persons receiving incomes; 28 per cent of the income is included in the area beyond the \$3000 line, and this is received by but 6 per cent of the persons receiving incomes; 40 per cent of the total income is included in the area beyond the \$2000 line, and this is received by but 14 per cent of the persons receiving incomes. It is clear that the distribution of incomes is very unequal. The height reached by the huge incomes is truly dazzling, viewed from the standpoint of the incomes received by the masses.

EQUALITARIAN DISTRIBUTION CONSIDERED

These facts regarding inequality of income are very impressive. It is no wonder that many persons, centering attention upon the enormous incomes of the rich, jump to the conclusion that if only there were more equal distribution of income everyone would be comparatively well-to-do. Let us experiment a little with the possibilities in this direction, having in mind our income chart. Suppose we chop off entirely all the incomes above \$4000 and add them pro rata to the incomes below \$4000, leaving only an average income for the 3 per cent deprived of their incomes. How much would be added by this leveling-down process to the moderate and lower-range incomes? Only 24 per cent. If we chop off all incomes above \$3000, leaving only an average income to the 6 per cent thus deprived, there would be added to the incomes below \$3000 a pro rata increase of only 30 per cent. If we level down the incomes above \$2000, leaving an average income for the 14 per cent who would thus be deprived, we should add 43 per cent to the lower incomes on a pro rata basis.

To be sure, an increase of 24, 30, or 43 per cent is not inconsiderable, but it is clear that the dreams of affluence of equalitarians or near-equalitarians would not be realized. As a matter of fact, an addition of 24, 30, or 43 per cent to the real incomes of the lower groups could not

possibly be realized; for out of the large incomes above the \$3000 and \$4000 range come the major proportion of the savings which go annually into new capital equipment as well as a large part of the government revenues. Some taxes have been deducted from these incomes, since the distribution is estimated in part from the Federal income-tax returns, which permit taxpayers to deduct "personal taxes and all taxes on property not used for business purposes, except special assessments to pay for improvements which benefit property."¹ But other taxes have not been deducted, and therefore the net incomes in the higher brackets are considerably smaller than here shown. In 1935 the sum of \$1,240,000,000 was paid in Federal income and inheritance taxes alone, and this, of course, was paid out of the higher incomes. In 1936, about 10 per cent of a \$12,000 income in New York State was taken by Federal and state income taxes, about 22 per cent of a \$40,000 income, 40 per cent of a \$100,000, and 75 per cent of a \$1,000,000 income. Moreover, we are safe in saying that the great bulk of the total savings which go into new capital equipment comes from the higher income levels. Independent estimates made by King for the United States and by Stamp and Pigou for the United Kingdom indicate that about 12 to 15 per cent of the national income is put aside in investment in new capital, including new houses, annually.² A considerable part of this comes, it is true, from corporate surpluses, which are not included in the income as here given. Even so, it is evident that a large part of the income of the more well-to-do goes to taxes and capital accumulation. Since these payments would have to be made by somebody, it follows that the addition which could be made to lower incomes by leveling down all incomes above \$2000, \$3000, or \$4000 would be much less than the 43, 30, or 24 per cent suggested above.

It is perfectly clear that there is not much to be gained by the leveling-down process. As a matter of fact, a far greater increase in real income has been achieved by the masses during the last hundred years through increased per capita production than could possibly come through an absolute equalization of incomes, even though we assume that such equalization would have no deleterious effect on national production. It should be noted, however, that if a considerable part of the incomes of the richest class were added to the incomes of the *poorest class alone*, a very material improvement in their status would be effected.

¹ W. C. Mitchell, *Income in the United States*, p. 139.

² In recent years a smaller per cent of the British income goes into savings and more into taxation. See A. C. Pigou and C. Clark, *The Economic Position of Great Britain*, 1936.

PRODUCTION VERSUS DISTRIBUTION

A hundred and thirty years ago two revolutions were sweeping over Europe. The one was a revolution in the methods of production; the other was a revolution in social structure. The first occupied the stage in England; the second, in France. The inventors and production engineers believed that the Industrial Revolution would speedily bring an abundance of material goods within the reach of all mankind at the expenditure of only a few hours of labor a day. The social idealists believed that the French Revolution, with its watchwords of "Liberty, Equality, and Fraternity," through its leveling process would uproot poverty. To the first group the chief cause of poverty was insufficient production; to the second group, it was maladjustments in social institutions, inequality of opportunity, and unequal distribution of wealth. Social reform is by no means unimportant, and it is as true today as formerly that eternal vigilance is the price of liberty, but it is none the less true that such gains in material well-being as the last hundred years have witnessed are chiefly the result of inventions and improvements in production. More than any other group it is the inventors and the production engineers who have revolutionized modern life and raised the level of well-being.

PERMANENCY OF RELATIVE DISTRIBUTION OF INCOME

In spite of vast changes in social institutions it appears that in Great Britain and probably in the United States the relative distribution of income has remained substantially the same for a hundred years. Sir Josiah Stamp, in *Wealth and Taxable Capacity*, gives data on the distribution of all incomes over £200 (roughly about \$1000) in 1801 and 1920. Table 43 gives this distribution.

Attention should be called first of all to the fact that this table excludes from consideration the great bulk of the population whose incomes fall below £200. Only 5.7 per cent of the income-receivers had incomes above £200 in 1920, and in 1801 the proportion was much smaller; therefore the table includes only the incomes of the relatively well-to-do. But if we take the table as it stands, it appears that a relatively larger proportion of all the incomes over £200 fell into the middle groups (£500 to £2000) in 1801 than in 1920, and that the proportion which fell into the lowest class (£200 to £500) was greater in 1920 than in 1801. Moreover, the richest 1.3 per cent received 24.2

TABLE 43. Distribution of Incomes in Great Britain in 1801 and 1920

PERIOD	INCOME CLASS	PERCENTAGE OF TOTAL NUMBER OF INCOMES OVER £200	PERCENTAGE OF TOTAL INCOMES OVER £200
1801	£200 to £500	61.5	24.0
	£500 to £1000	21.3	18.6
	£1000 to £2000	10.3	17.7
	£2000 to £5000	5.3	20.3
	Over £5000	1.4	19.5
1920	£200 to £500	71.3	29.0
	£500 to £1000	15.8	15.8
	£1000 to £2000	7.8	15.4
	£2000 to £5000	3.7	15.4
	Over £5000	1.3	24.2

TABLE 44. Distribution of All Incomes over £200

INCOME-RECEIVING CLASS	PERCENTAGE OF TOTAL INCOMES OVER £200 RECEIVED BY EACH CLASS	
	1801	1920
Richest class, constituting 2 per cent of all incomes over £200	22.5	28.5
Next richest class, constituting 8 per cent of all incomes over £200	25.0	22.5
Rich class, constituting 15 per cent of all incomes over £200	18.5	17.0
Well-to-do class, constituting 25 per cent of all incomes over £200	16.0	14.5
Lower half, 50 per cent of all incomes over £200	18.0	17.5

per cent of the income in 1920, whereas the richest 1.4 per cent obtained only 19.5 per cent in 1801. This would indicate that income was somewhat less highly concentrated in 1801 than in 1920. This fact is brought out more clearly in Table 44, which shows the distribution of income as we have estimated it from Stamp's data for the richest 2 per cent, the next richest 8 per cent, the rich 15 per cent, the well-to-do 25 per cent, and the lower 50 per cent of all incomes over £200. From this table it is evident that the richest class received relatively more of the income in 1920 than in 1801. The difference, however, is not great.

Stamp calls attention to the fact that incomes were not as fully reported in 1801 as in 1920, and that possibly there was relatively greater evasion in the lowest-income group, the one between £200 and

TABLE 45

PERIOD	INCOME CLASS	PERCENTAGE OF TOTAL NUMBER OF INCOMES OVER £500	PERCENTAGE OF TOTAL INCOMES OVER £500
1801	£500 to £1000	56.0	24.4
	£1000 to £2000	26.3	23.2
	£2000 to £5000	13.9	26.7
	Over £5000	3.8	25.7
1920	£500 to £1000	55.2	22.3
	£1000 to £2000	27.3	21.8
	£2000 to £5000	13.0	21.8
	Over £5000	4.5	31.4

TABLE 46. Distribution of All Incomes over £500

INCOME-RECEIVING CLASS	PERCENTAGE OF TOTAL INCOMES OVER £500 RECEIVED BY EACH CLASS	
	1801	1920
Richest class, constituting 5 per cent of all incomes over £500	29.5	35.5
Next richest class, constituting 15 per cent of all incomes over £500	26.0	22.5
Rich class, constituting 30 per cent of all incomes over £500	23.5	20.5
Lower half, 50 per cent of all incomes over £500	21.0	21.5

£500. If we exclude this group, we have for incomes over £500 the distribution shown in Table 45.

From Table 45 it will be seen that the distribution of incomes above £500 was closely similar in the two periods, but again it is evident that the richest class received a larger percentage of the total income in 1920 than in 1801. This fact can be seen more clearly if we estimate the proportion of the total income received by the richest 5 per cent, the next richest 15 per cent, the rich 30 per cent, and the lower 50 per cent of all incomes over £500. These data are given in Table 46. Here again it appears, as in Table 44, that the relative gain of the richest class has come chiefly at the expense of the next richest classes and not at the expense of the lowest-income group.

While the relative distribution of incomes appears to be substantially similar in 1801 and in 1920, the *proportion of the population* included in the income levels given was very much greater in 1920 than

in 1801. Stamp reaches the following conclusion: "The total nominal income has increased much more than the total population—the increase has surged upwards through all the fixed classes, so that there is a smaller population in the ranks of the poorest, with a nominal income of, say, under £80 a year, and many more in the over £5000 class, but the *slope* of the distribution, that is, the relation between one section or class and another, has hardly altered."

There is thus an astonishing similarity in the distribution of income in two periods separated by a century and a quarter, during which time revolutionary changes had occurred in economic and political institutions. Similarly, Professor A. L. Bowley in *The Change in the Distribution of the National Income, 1880-1913*, concludes that the increase in the per capita income (about 35 per cent) in the period in question was shared with remarkable equality among all the various classes, including the poorest wage-earners. This suggests the possibility that the dominant forces controlling the distribution of income hitherto have been deeper and more fundamental than mere social institutions. It would seem to point to a "fixed system of causation" such as the great inequality of native capacities in the human race. This explanation overlooks the important fact, however, that throughout the period in question the basic institutions of private property and the inheritance of private property have prevailed intact. The permanence of these fundamental institutions, in addition to the relatively fixed distribution of inherited capacities, probably accounts in large part for the persistence of the scheme of income distribution which has prevailed in modern communities during the last century or more.

Tables 47 and 48 give data relative to income distribution in the United States in 1926 and in Great Britain in 1929.

That the distribution of income does not conform to the distribution of inherited capacities is clear from an examination of Fig. 42. The income-distribution curve, if enlarged to include all incomes, would be lopsided to an extreme degree; on the other hand, it is likely that a curve showing the distribution of ability would be fairly symmetrical. Above all, it is the institutions of private property and the inheritance of private property which distort the income curve. Indeed, the lower part of the income curve, between zero and \$2000, is fairly symmetrical, and these incomes are derived largely from personal services and only to a small extent from property.

Yet it is probable that even though all income were earned from personal services, the income curve would still be skewed to the right,

TABLE 47. Income Distribution in United States in 1926¹

INCOME CLASS	NUMBER OF INCOME RECIPIENTS (IN THOUSANDS)	PERCENTAGE OF TOTAL INCOME RECEIVED
Under \$8,309	44,195.0	87.277
\$8,309-\$41,543	429.0	7.275
\$41,543-\$249,255	47.0	3.792
Over \$249,255	2.5	1.656

TABLE 48. Income Distribution in Great Britain in 1929²

INCOME CLASS	NUMBER OF INCOME RECIPIENTS (IN THOUSANDS)	TOTAL INCOME IN CLASS (IN MILLIONS OF POUNDS STERLING)
Under £125	11,600	1,170
£125-£250	4,925	980
£250-£500	1,527	404
£500-£1000	508	312
£1,000-£2,000	199	237
£2,000-£10,000	100	378
Over £10,000	10	221

though certainly much less so than under existing property institutions. The artist who is just a little better than his nearest rival can command a much larger price for his paintings. The physician who makes a great reputation may be barely, if at all, superior to his competitors, but he will be able to demand much larger fees for his services. The situation appears not altogether unlike a race in which the winner obtains a large prize and his nearest competitor receives nothing at all.

PROPERTY INSTITUTIONS AND SOCIAL UTILITY

In so far as inequality in income distribution is due to the inheritance of private property, no justification for such inequality can be found on the basis of personal merits and demerits.³ The person who inherits property, such as natural resources or capital goods, receives an income for the services of his property. The reward is for service rendered, but it is a *property service* and not a *personal service*.

¹ W. I. King, *The National Income and its Purchasing Power*, 1930.

² Pigou and Clark, *The Economic Position of Great Britain*, 1936.

³ An important exception to this statement is found in those cases in which the children have worked in their father's business without adequate financial return in the form of wages, educational opportunities, or other reward for the services rendered.

How, then, can we justify private property, especially the inheritance of private property, and the inequalities to which it gives rise? They can be justified only on the ground of *social utility*. It can scarcely be questioned that inequality of income has facilitated in a marked degree the rapid accumulation of capital which we have witnessed during the last century. This does not mean, however, that a socialist society cannot also accumulate capital effectively; the Russian experiment demonstrates that. But it is true at any rate that the private accumulation of capital has contributed greatly to such economic progress as we have made during this period. Without it we could never have reached the per capita income which we have attained. Inequality of income thus has been one of the factors that have enabled us to raise the *level of income* all round, but it has left us with the same degree of inequality. Other aspects of the social utility of private property will be discussed in Chapter XXXI ("Socialism").

With the wider diffusion in the ownership of stocks, bonds, savings deposits, and life-insurance policies among the wage-earners, must we not conclude that property and income are becoming more equally distributed? Not at all, since the middle groups and the rich also are accumulating more. But must we not at least conclude that with the continued increase in per capita income, capital accumulation would not be slowed down even though income were more equally distributed? The answer is not certain. There are signs that human wants and desires are increasing even faster than wealth and income, and it is by no means clear that it is easier now than formerly to save. If indeed the increase in average income should result in an increase in the per cent of the national income which is saved, a lower interest rate would tend to prevail and this in itself would tend toward greater equality in income distribution. Thus inequality of income and capital accumulation interact upon each other.

FUNCTIONAL DISTRIBUTION

We turn now to a brief consideration of the statistical data relating to the distribution of income from the standpoint of the factors of production. According to Dr. W. I. King, the national income was distributed in the four functional shares approximately as follows in 1910:

Wages and salaries	46.9 per cent
Interest	16.8 per cent
Rent	8.8 per cent
Profits	27.5 per cent

Inequalities of Property-Ownership and Income 559

Dr. King, however, includes under profits the wage or salary that should be imputed to the individual entrepreneur—businessman or farmer. He estimates roughly that three fourths of the profits so defined are really wages and salaries. If we add this imputed item to wages and salaries we have the following distribution :

Wages and salaries	67.5 per cent
Interest	16.8 per cent
Rent	8.8 per cent
Profits	6.9 per cent

Similarly, the National Bureau of Economic Research estimated that in the ten-year period from 1909 to 1918 approximately 71 per cent of the net-value product of mines, factories, and railroads went to wages and salaries, leaving about 29 per cent for profits, interest, and rent.

For the year 1928, the same authority estimated that 57 per cent of the total income of the people of the United States was received as wages and salaries, and 43 per cent as the share of property-owners and of entrepreneurs. It should be noted that the latter includes the income of all farmers and small businessmen, and that a large part of their income is a reward for their own labor and this part ought really to be added to the "wages and salaries" category. When this is done the figure for 1928 will compare rather closely with those for 1910 and 1909-1918.

With respect to the different industries the national income was derived as follows in the years given :

TABLE 49

	1913 <i>(in Millions of Dollars)</i>	1918 <i>(in Millions of Dollars)</i>	1928 <i>(in Millions of Dollars)</i>
Agriculture	5,887	12,682	8,109
Mines, quarries, and oil wells	1,191	2,013	2,169 ¹
Factories	7,976	16,018	16,866 ¹
Construction	1,669	1,280	3,458 ¹
Other hand trades	852	1,704	
Railroads	2,181	3,570	4,085 ²
Other public utilities	1,154	1,661	2,783 ¹
Banking	509	767	1,165 ²
Government (all branches of government in the United States)	1,829	5,352	6,130
Unclassified industries and miscellaneous income	12,332	15,318	36,388

¹For the year 1925. See W. I. King, *The National Income and its Purchasing Power*, pp. 94, 95.

²For the year 1926. See W. I. King, *ibid.*

FUNCTIONAL AND PERSONAL DISTRIBUTION CONTRASTED

Functional distribution deals with the division of the social product among the four factors of production. It explains how the prices of the services of these factors are determined, but it does not explain the manner in which the earnings of these factors are distributed among the people who make up the community. For example, a businessman may receive an income of \$20,000 annually from his business, a composite of wages, interest, and pure profits. The theory of functional distribution explains why he receives that amount, but it does not tell us what he will do with it. A part, we know, will be taken by taxation and used to support various governmental services in which the people of the community share according to the legislation regulating governmental expenditures. Another part will be consumed by the man and his family, and within the family group the distribution in the main is according to need. Still another part may be donated to the support of charitable institutions, whose benefits also will be shared according to the principle of need. The ultimate distribution of the annual real income of society is not made according to the principles of value and distribution, because those principles do not control the disposition of money incomes after these incomes have passed into the hands of the owners of the factors of production; they do, however, control the amount of money income the owners of these factors receive.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The number of stockholders in the United States has increased considerably, it is said, since 1920. Should you conclude from this that ownership of property is less highly concentrated than formerly? Why, or why not?
2. In the United States 3 per cent of the population receive about 22 per cent of the total income, whereas in the United Kingdom the same proportion of the population receive about 33 per cent of the income. How do you account for this difference?
3. Inequality in the distribution of income is not nearly so great as inequality in the distribution of property. Why is this the case?
4. Rank in order of importance the following causes of inequality in the distribution of wealth and income: inheritance of private property, unequal educational and social opportunities, differences in native capacity, chance gains, monopoly power.
5. "Let me emphasize again what I have said before, that it is probably in and through the exercise of the principle of plunder or the undue exercise of ad-

vantage, of gambling or of its allied principle of monopoly, or of special privilege or favor of some kind that many, very many, if not most of the greatest fortunes have been won.”—J. W. JENKS. Evaluate this statement.

6. According to Professor Sorokin, in a study entitled “American Millionaires” (*Journal of Social Forces*, May, 1925), 39 per cent of deceased American millionaires started life poor and 30 per cent started life rich. Of living American millionaires only 20 per cent started life poor and 53 per cent started life rich. What does this difference indicate?

7. Professor Sorokin finds that of the wealthy men who started their careers poor, 13 per cent had received no formal education, 58 per cent had an elementary education, 17 per cent a high-school education, and only 12 per cent a college education. Of the wealthy men who started life in moderate circumstances, 2 per cent had no education, 28 per cent had elementary education, 27 per cent high-school education, and 42 per cent college education. Of those who started life rich, 7 per cent had only an elementary education, 13 per cent had a high-school education, and 80 per cent a college education. All told, 54 per cent of the American millionaires held college degrees.

Of those who started their career poor, only 2 per cent became rich before the age of 30, 23 per cent between the ages of 31 and 40, 39 per cent from 41 to 50, and 36 per cent above 50; of those who started life rich, 27 per cent became responsible managers of big business before reaching the age of 30, and 53 per cent between the ages of 31 and 40.

8. The following is taken from the *Minneapolis Journal* for July 27, 1924: “In St. Paul, Probate Judge Howard Wheeler issued a statement showing that 60 to 70 per cent of all adults dying in Ramsey County last year left no more estate than enough to pay their funeral expenses; that although 2353 adults died, only 860 estates were probated; that of that number only 25 left more than \$50,000, 90 from \$10,000 to \$50,000, 121 from \$5000 to \$10,000, and 20 left an estate of no definite value or consisting only of a cause of action for wrongful death.”

CHAPTER XXXI · Socialism



THE EXTENT OF THE SOCIALIST MOVEMENT

The fundamental purpose of this textbook is to explain as fully as possible within the limits of one volume the economic phenomena of our present social order. Yet we cannot ignore the fact that the economic order is a changing one. Nor is it possible to overlook the fact that a considerable portion of the voters of all modern countries are seriously dissatisfied with the present order and desire a fundamental reconstruction of our economic system. Out of a total of 30,000,000 votes cast in the 1919 election in Germany, about 45 per cent were socialist and communist votes. In the 1924 election the combined socialist and communist vote fell to about 35 per cent of the total; but in the election of 1928 it increased again to about 40 per cent of the total. In the 1932 election, 38 per cent of the seats in the Reichstag went to the socialists and communists combined. The advent of Hitler to power, however, has either liquidated or driven underground the socialist and communist movements. In the 1924 election in Great Britain, 5,552,000 votes out of a total of 16,564,000 were cast for the Labor party, which has a socialist program. Thus exactly one third of the British electorate voted the labor ticket. In the 1935 election, 42 per cent of the vote went to the Labor party. In France the socialists elected 106 deputies to the lower house in the election of May, 1924, and the communists elected 29. In 1936 the socialists elected 145 deputies and the communists 71, about 35 per cent of the total. In the smaller countries of northern and western Europe, about one third of the electorate vote the socialist ticket. Moreover, in Russia we have seen a stupendous experiment in state communism. Not only do socialist representatives constitute about one third of all the representatives in the popular branch of the legislative bodies in northern and western Europe, but in nearly all these countries socialists have held important positions in the cabinets, and four countries—Sweden, Great Britain, Denmark, and France—have had socialist governments. Many municipal governments, moreover, are controlled by the socialists. In the United States socialism as a political

movement is weak in comparison. A few cities have had socialist mayors, notably Milwaukee, Minneapolis, and Schenectady. In 1920, of a total Presidential vote of 26,700,000, the socialist vote was only 920,000. In the 1936 election, out of 45,800,000 votes, fewer than 200,000 went to the socialist candidate.

ORIGIN AND DEVELOPMENT OF SOCIALISM

This extensive and widespread movement is a newcomer among political parties. As a political movement it started in France under the leadership of Louis Blanc about 1840. For a short time the socialists shared in the provisional government established after the revolution of 1848. Fifteen years later a socialist political party was formed in Germany under the leadership of Lassalle. This united with another socialist group, followers of the great socialist Karl Marx, in 1875. This party, known as the Social Democratic party, made rapid strides in Germany, and by 1890 polled 20 per cent of the total vote cast. By 1912 this was increased to 35 per cent. During the war the party split into three groups, and the movement subsequently divided into the socialist and communist parties. In England small socialist parties were organized in the eighties and nineties, but it was not until 1906 that the British Labor party, which soon adopted a socialist program, was definitely organized. Its growth was rapid, and it soon won an important representation in Parliament. In France the socialists participated for a brief time in the provisional government of 1871, but the excesses of the Commune and the subsequent banishment from the country of leading communards completely disrupted the movement. In the eighties five different socialist parties emerged, some moderate and others revolutionary. By 1900 kindred groups had merged sufficiently so that the number was reduced to two, one revolutionary in its views and the other moderate. These united in 1905 and formed the Unified Socialist party. In 1936, M. Blum became the first socialist prime minister of France.

UTOPIAN SOCIALISM

Modern political socialism was preceded by a nonpolitical group of socialists, frequently termed utopians. The outstanding leaders of this group were Robert Owen in England and Charles Fourier in France. Both wished to establish voluntary communities, each to consist of from two thousand to four thousand members. Owen founded, in fact, a colony

in New Harmony, Indiana, in 1825, and scores of Fourieristic communities were established in the United States in the forties. None of these colonies consisted of more than a few hundred members. All of them failed in the end, but several survived for a number of years. Owen's community was communistic, all property being held in common, the proceeds being shared equally by all. Fourier's communities were merely associative: the property of the community was owned by a stock company, the shares of which were subscribed for by such members as had capital, and these shareholders were to receive a certain percentage of the surplus production, the rest being distributed according to labor and ability; production and consumption were to be on an associative basis, but there was to be no equal distribution of income, nor was private property to be dispensed with.

Both Owen and Fourier made their appeal to all classes of society, but more especially to the powerful and the well-to-do. They believed that the existing order was all wrong and that this unfortunate condition was due simply to human ignorance. No one heretofore had set forth the true and ideal social order. They thought they had discovered such an order and believed that it was necessary only to inform people about it. Rational human beings would see the reasonableness and superiority of the new order, and all groups and classes would adopt it. Such, in brief, was the essential point of view of the utopian socialists.

MARXIAN SOCIALISM

In 1848 Karl Marx and Frederick Engels wrote the *Communist Manifesto*, which presented in brief but brilliant fashion the point of view of what they conceived to be a "scientific socialism." The *Manifesto* did not set forth an ideal order which people ought to adopt. Instead it was an interpretation of the evolution of economic society. It sought to explain what was going on in the economic order and to predict on the basis of current tendencies what the future would bring. The central doctrine is the "materialist interpretation of history," as Engels called it. All social and political history is but a record of the conflict of economic classes. An economic class has common interests because its members have the same occupational and income status. But its interests come into conflict with the interests of other economic groups; thus, the economic interests of landlords are not the same as the economic interests of renters, nor are the interests of capitalists the same as those of wage-earners. Often what benefits one group injures

another. Thus it is argued that wherever society is divided into economic classes a conflict of interests develops, and the social and political movements to which this conflict gives rise constitute history. Now, according to Marx and Engels, the alignment of economic classes is a product of the prevailing mode of production and exchange.¹ Thus it was the widening of the market and the development of the domestic system of manufacturing and finally large-scale production that created the modern capitalist class on the one hand and the wage-earning class on the other. In brief, changes in the technique of production create new economic classes, these class alignments give rise to a conflict of interests, and this conflict of interests results in the social and political movements which make up historical events.

In the *Communist Manifesto*, Marx and Engels sketched the historical rise of the modern capitalist class. They pointed out how the rise of this new class inevitably resulted in the downfall of the old feudal monarchies. The rise of this class into a position of pre-eminence found expression politically in the establishment of modern political democracies. Representative government based upon the suffrages of the middle or capitalist classes wrested political power and special privileges from the landed nobility.

But capitalism, according to Marx and Engels, cannot grow and develop without at the same time pushing a larger and larger portion of the population into the ranks of wage-earners. Large-scale production more and more eliminates the middle classes. Thus the development of capitalism inevitably gives rise to a powerful working-class movement. The working class will finally rise to a position of pre-eminent power and will thus control not only the political state but also industry itself. The capitalists will be ousted from control, all means of production will be socially owned, and government and industry will be controlled by and carried on in the interest of the working class.

Marx and Engels thus predicted that socialism was inevitable. This prediction was based on two fundamental postulates: first, that the middle class would inevitably be wiped out by the development of large-scale production; second, that the working class would inevitably sink into deeper and deeper misery by the very nature of the capitalist system. If these two propositions were indeed true, a revolutionary upheaval no doubt would be inevitable. Whether or not a socialist society would then follow would depend upon its workability.

¹See Alvin H. Hansen, "The Technological Interpretation of History," *Quarterly Journal of Economics*, November, 1921.

As a matter of fact, however, capitalism hitherto has not eliminated the middle class, nor has it resulted in the progressive deterioration of the working class. It is quite true that the precapitalist type of middle class is declining relatively. The *proportion* of our population who are individual proprietors is decreasing. Yet the total number of individual proprietors is not declining: farming still persists as an individual-proprietorship business, and there still continue to be an enormous number of small merchants, bankers, and even manufacturers. In any large city one cannot fail to notice many small manufacturing plants. The old type of middle class thus continues to survive in large numbers. But, more than that, the capitalist system has developed a new type of middle class peculiar to itself. The great body of technical experts, managers and superintendents, scientists, and professional people, without which the highly complex modern industrial society could not possibly operate, constitute a new capitalist middle class.¹ Not only from the standpoint of income and social status do they constitute a middle class, but also, in large part, from the standpoint of property. The property owned by the new middle classes is not of the individual-proprietorship type: it consists largely of stocks and bonds, mortgages, insurance policies, and savings accounts. Many recent researches indicate that the middle classes are not becoming appreciably poorer relatively to the rest of the population, either as to property or as to income. It is true that in Germany a considerable proportion of the middle class was wiped out by the uncontrolled price inflation following the war, but this was in no sense a normal development of capitalism.

Marx and Engels were quite right in assuming that the middle class constitutes the great bulwark of the present economic order; they were quite right in assuming that if this class were wiped out, capitalism would receive a body blow which it might not survive; but so long as the middle class persists and finds its economic interests rooted in the present economic order, capitalism will be difficult to overthrow.

Nor has the development of capitalism resulted in the impoverishment of the working classes; on the contrary, real wages more than doubled from 1850 to 1900, and were, until the recent depression, higher in leading industrial countries than ever before. The current revival is restoring this level, and in some cases a new high level of real wages for employed workers has been established. Marx believed that

¹See Alvin Hansen, "Industrial Classes in the United States in 1920," *Journal of the American Statistical Association*, December, 1922; and Tillman M. Sogge, "Industrial Classes in the United States in 1930," *ibid.*, June, 1933.

under capitalism new machinery was being introduced so rapidly that it necessarily supplanted labor, creating a larger and larger reserve army of unemployed, thus progressively lowering wages. As a matter of fact, however, the increased output consequent upon the introduction of machinery gave rise to an increased demand for products; and the increased profits due to the cheaper machine processes stimulated the accumulation of capital and resulted in the building and operation of numerous new plants which reabsorbed the labor supply. Thus machinery in the end has not displaced labor, but instead has increased the marginal product of labor and thus increased real wages. From the long-run standpoint there is no convincing evidence that machinery is creating a permanent unemployment problem. The great problem of unemployment is caused not by machinery, but by the business cycle, by world-wide structural changes such as those which followed in the wake of the World War, and by increasing rigidity of prices and wages.

Nevertheless improved conditions do not in themselves prevent the labor movement from becoming socialistic. Labor is eager, as are all groups in society, to get all that it can for itself, and in so far as the working class becomes convinced that socialism will offer more than capitalism, it will naturally become socialistic. Nevertheless a *revolutionary* overthrow of capitalism is not likely in a democratic society in which the conditions of the great majority of wage-earners are improving. It was a *violent* overthrow of capitalism that Marx and Engels predicted.

EVOLUTIONARY SOCIALISM

By the last decade of the nineteenth century many of the keener minds in the socialist movement had become thoroughly aware that the tendencies pointed out by Marx were not working out as predicted. Moreover, the socialist movement was gaining in political power and influence, and it was perforce being pushed into a practical program of social reform along such lines as labor legislation, social insurance, and taxation. Thus both the theory and the practice of socialism were becoming evolutionary and opportunistic. In the eighties the Fabian socialists in England challenged the doctrines of Marx and sponsored a gradual reconstruction of society, which, however, it was expected, would lead toward the socialist goal. In 1894 the Austrian economist Böhm-Bawerk wrote a crushing criticism of Marx's theory of value and distribution. Since Böhm-Bawerk's attack no one has ever been able to reconstruct the Marxian system of theoretical economics. Finally

in 1899 Bernstein, a German socialist writer, admitted the collapse not only of the Marxian economic theories but also of Marx's predictions as to the tendencies of capitalism. Jaurès, the outstanding figure of the French socialist movement before the World War, stood for an evolutionary program of social reform. In his view socialism would come piecemeal, so gradually that no one would be able to say when we had passed from capitalism to socialism. Thus political socialism ceased to look forward to a final revolutionary overthrow of capitalism. It buckled down to a program of social reform involving increasing labor legislation and more government ownership. This, it was believed, would finally lead to socialism.

SYNDICALISM

In the meantime, however, there developed a new revolutionary movement, syndicalism. It had its origin in France in the decade of the nineties. Its essential philosophy contained elements derived in part from anarchism, in part from socialism, and in part from trade-unionism. Unlike the socialist movement, it was antipolitical in its methods, and placed its reliance upon sporadic strikes, the general strike, and sabotage (a method which involves staying on the job but interfering by slovenly work or destructive acts with the profits of the employer, thus forcing him to come to terms). Unlike the current trade-union movement, however, it did not lay the emphasis on immediate ends, such as wages, hours, and conditions of work, but instead on revolutionary ends, a radical reconstruction of society. A great general strike was finally to bring about a collapse of the present system. In its place there was to arise a new society which was to consist of a general federation of *syndicats*, or trade-unions. The local unions were to be federated into city federations; local unions in different parts of the country belonging to the same trade or industry were to be federated into national unions; and finally these national unions and the city federations were to be federated into a General Federation of Labor. This in part was the type of organization already existing in the French General Confederation of Labor. Thus the existing trade-unions were to be the growing cells from which would spring the new social order. The structure was already developed, ready to assume control of the new order. Instead of all industry's being controlled by a bureaucratic socialist state, each industry would be controlled and managed by the *syndicat* or *syndicats* covering that field of work. But the ownership

would inhere in local federations of *syndicats* in the case of local industries and in national *syndicats* for national industries.

Syndicalism is thus similar to anarchism in that its social structure is a loose decentralized federation of more or less independent units. But the basic unit of syndicalism is the trade-union. All society is thus organized from the standpoint of the producer. With anarchism, on the other hand, there would be federations (purely voluntary) for all sorts of purposes: federations for agricultural and industrial production, and federations for consumption that would make provision for food, fuel, electricity, and dwellings.

GUILD SOCIALISM

During the war and postwar years guild socialism developed in Great Britain. It was a compromise between syndicalism and political socialism. It aimed to retain the present state, which would own all the means of production. But industrial guilds (including all manual workers, technicians, and managers) would manage the several industries; thus the coal-miners' guild would control the coal industry, the railway guild would control the railroads, and so on. Prices would be controlled by the state, and industry would be taxed by the state for its own maintenance and the provision of public services. The proceeds remaining after the payment of this tax might be distributed by the guild to its members on any basis it saw fit. Another school of guild socialists would substitute the "commune" for the state; this would be composed of representatives from civic organizations and from consumers' and producers' guilds.

COMMUNISM

The Russian revolution has brought to the front another socialist group, who call themselves Communists to distinguish themselves from the moderate political socialists. (Before the Russian Communists established themselves in power they were known as Bolsheviks, *bolshevik* meaning "of the majority." At the time the name was given they were the majority in the Russian Social Democratic party.) Their methods were revolutionary rather than evolutionary in character. They believed that socialism could be introduced at once and did not need to wait, even in such backward countries as Russia, for the complete development of capitalism. They believed that socialism could not be established without disfranchising employers and prop-

erty-owners and introducing a "dictatorship of the proletariat," or propertyless masses. They believed in the expropriation of property by the state without compensation to private owners. They had no use for the democratic political state, but wished to establish in its place a "soviet republic." The soviet state was to consist of "soviets," or councils of workingmen in the urban districts and peasants in the rural districts. These local soviets in turn were to be represented in provincial and regional soviets, and these in turn in the all-Russian Congress of Soviets.

The Russian Communist party, in spite of the fact that it had a membership of only some 600,000, succeeded in retaining control of the Soviet government after the autumn of 1917. This remarkable feat was accomplished by the following methods: (1) voting by show of hands instead of by secret ballot, (2) disqualification of opposition candidates on some pretext or other, (3) repression of opponents' propaganda, (4) Communist control of press and schools, (5) ousting of non-Communist trade-union officials and substitution of Communists, (6) ousting of non-Communist officials of co-operative societies and placing of Communists in control of the central organization, (7) officering of the army by Communists, (8) opening of military schools only to Communists, and (9) the Terror, by which internal foes of the government (particularly those in sympathy with foreign intervention) were ruthlessly suppressed.

In November, 1936, a new constitution was adopted, under which the former Congress of Soviets was supplanted by a bicameral parliament similar in form to that of Western democracies. The new constitution represents a formal change, but not necessarily any movement of the Russian dictatorship toward a more liberal and democratic system in which the people will have more civil rights and more initiative in the government. Shortly after the inauguration of the new constitution, an alleged conspiracy against the government led to the execution of nearly a score of outstanding persons, all former leaders in the Communist government.

The Russian Communists immediately upon coming into power proceeded to nationalize the land, but the peasants were given "possession" of it for use, subject only to a levy on their surplus products. Next the Communists proceeded to nationalize the factories, the mines, and the banks. By 1919 the process of nationalization was practically complete. But the extreme communism of this first period fell down on three counts: (1) it failed to maintain such capital equipment as existed,

and it was utterly unable, except in sporadic cases, to make provision for new capital equipment; (2) it failed to secure the effective cooperation of the former owners, managers, and technical experts; (3) it failed to motivate the working class to maintain the former level of personal efficiency. Moreover, the Communists were unable to import much machinery and railroad equipment from abroad, at first because they were engaged in a war with foreign countries on several fronts, and later because of the reluctance of foreign governments and capitalists to enter into trade relations with them. So long as the government was engaged in defending the revolution against foreign powers the peasant classes, constituting the overwhelming majority of the population, acquiesced measurably well in the Communist rule; but as soon as the menace of foreign intervention disappeared the peasants became restless. Under the land law they were compelled to turn all their surplus products over to the state. This policy discouraged production on the one hand and resulted in secret barter on the other. The peasants demanded a repeal of the law conscripting all their surplus, and freedom of trade so that they could dispose of their products. This the Communists were compelled to grant for a time. Moreover, they granted the right to lease land, to hire laborers, and to inherit a limited amount of property. Furthermore, the breakdown of factory production was so great that to get needed capital and increase production they were compelled to permit for a time a partial return of private enterprise except in transportation, large industry, and export trade.¹ Concessions were granted also to foreign capitalists. Thus private enterprise and freedom of trade to some extent again were introduced. Private enterprise returned in some measure in agriculture, in small-scale manufacturing, and to a considerable extent in domestic trade; but the bulk of the large industries were organized in great state trusts. The "New Economic Policy," with its partial abandonment of communism, resulted in marked improvement in production.

This New Economic Policy continued in force from 1921 to 1928. In the latter year began an era of economic planning and a marked trend toward increased collectivization.

A "Five-Year Plan" was inaugurated, designed to hasten the industrialization of Russia. Special attention was devoted to the modernization of the heavy industries, coal, petroleum, electrical energy, blast

¹The leading remaining government industries were "transportation, textile, metal, mining, electrical, fuel, lumber, chemical, and construction." See Harry W. Laidler, *A History of Socialist Thought*, p. 503.

furnaces, machinery, and tractors. A second Five-Year Plan was launched in 1933, which laid more emphasis on light manufacturing industries and the production of consumers' goods.

In 1929 the government instituted a new agrarian policy, the collectivization of farms. Land, implements, and draft animals were placed under the control of collective farms, and these were supplied as far as possible with tractors and other machinery which could not be used effectively on small farms. Many peasants resisted collectivization and even slaughtered their animals rather than submit to the new policy. The program nevertheless was carried through, though it entailed a good deal of suffering. In 1933 a compromise solution was reached, under which the government collected a fixed grain tax from the collectives. The remaining surplus was left to be distributed by the collectives to the individual members in proportion to the work contributed by each. Each member could then sell or barter such grain in exchange for manufactured goods. Individual members were allowed to raise their own chickens and vegetables, and these could similarly be sold in the market. This program represents a compromise between individual farming and collectivization.

From 1929 to 1935 an elaborate system of rationing by cards was introduced, designed to limit the demand in conformity with the deficient supply of most consumers' goods. These rations permitted a worker to buy in the co-operative or "closed" shops the amount specified, at a price much below the free market price. If one wanted, and had the means, to buy a commodity in excess of the allotted ration, he would have to go to the "state commercial shops" where prices were 5, 10, and in some cases even 30 times higher than in the co-operative shops. Under the second Five-Year Plan, however, consumers' goods gradually became more plentiful and the margin between the "rationed" and "free" prices narrowed. By the fall of 1935 the ration cards had been abolished and "free" prices, somewhat higher than the rationed prices, but lower than the former "free" prices, became the usual rule in all shops.

Despite the higher degree of socialization of factories, farms, trade, and finance, it is still possible for an individual to acquire property in the form of a house, a car, and a few animals; or one can acquire state bonds; or if one prefers one can accumulate a deposit account in a savings bank whose funds are invested in state bonds. Property of this sort can be inherited up to an amount of 50,000 rubles. It is not possible, however, for an individual to own the means of production. Capital equipment is owned by the state, and business enterprise is carried on as

a state function. Nor can an individual live on an unearned income. While he may invest in state bonds which pay interest, incomes from such sources are not large, and in any event are heavily taxed. In general, income is received only as a return for work. A program of social security is designed to care for those who for various reasons are unable to work.

SOCIALISM AND COMMUNISM CONTRASTED

The democratic socialist parties of northern and western Europe which belong to the Socialist International are by no means in agreement with the Russian Communists; they differ from them on the following specific points: (1) The socialists believe in political democracy, in universal suffrage, and in the rule of majorities, they therefore are opposed to the dictatorship of a small militant minority such as that established in Russia by the Communists. (2) The socialists favor representative government which is based on geographical representation; the Communists favor the soviet system of government, though the recent constitution indicates a modification of their views on this point. (3) The socialists do not believe that it is feasible to nationalize all industries at once; they believe that the public administration of industry is such a difficult and complex undertaking that it must be entered upon piecemeal. (4) The socialists do not believe in expropriation of private property without compensation. As an industry becomes nationalized or municipalized they expect to make full compensation to the private owners, perhaps through the issue of government or municipal bonds. In the end, however, private property in such investments would gradually become less important under the continuous pressure of inheritance and income taxes. (5) The socialists do not believe it desirable to eliminate completely the freedom to engage in private enterprise. Only in the case of the great industries and municipal utilities, which are naturally monopolistic, do they intend to exclude private enterprise. They do, however, wish public enterprises or consumers' co-operative societies to engage in all fields and freely compete with private enterprise, and they believe that in such a situation private enterprise in the end will become relatively unimportant.

There is much difference among socialists as to whether there should or should not be complete equality in the distribution of income. There are many moderate socialists who believe that it is necessary to have unequal rewards in order to secure the effective co-operation of persons with superior ability. In actual fact, since the socialists would under-

take to socialize industry piecemeal, they would be compelled in the transition stage to adopt unequal rewards. Nevertheless the tendency in socialist theory is toward a system of equality. In this connection it is interesting to note that Russian communism has found it necessary to differentiate wage and salary payments according to performance.

SOCIALISM AND CAPITALISM

In view of what has been said above, it is clear that Russian communism is undergoing rapid change, and it is equally clear that capitalist societies are experiencing profound modifications in their economic structures. The development of authoritarian states in Italy and Germany represents an assumption of state control over economic life in many respects comparable with that of Russian communism, and even in the democratic countries it is evident that the capitalist order is becoming more and more subject to social control, and that the distribution of wealth and income is being profoundly modified through the application of progressive income and inheritance taxes.

The changes in taxation and social control that have taken place in the last few decades go far to justify the description of the present order as a socialized capitalism. Yet it is still a capitalist order so long as private ownership and private management are the rule. Whether or not the function of government will continue to expand until the social order may properly be described as a socialist one, it is impossible for anyone to predict. At any rate it is clear that there is a vast difference between a system of private property and private enterprise subject to extensive social control and a system in which public property and public enterprise is the rule. It is one thing to expect a certain amount of social control in private enterprise and property, and it is quite a different thing to undertake actively the public administration of all the hundreds of thousands of factories, shops, and farms. The program of socialization is by no means an easy one. Some of the most fundamental difficulties confronting a democratic socialism, in contrast to a dictatorship such as that in Russia, may be outlined for brief consideration.

DIFFICULTIES FACING SOCIALISM

1. Will a democratic socialist society be able to accumulate capital as efficiently as the present order? Will it be able to make adequate provision not only for the replacement of worn-out equipment but for

additional capital? Without replacement the entire capital equipment of a country would be exhausted in a very few years. The capitalist system has demonstrated over the past one hundred and fifty years its capacity to accumulate capital. Whether or not a democratic socialism could do as well is doubtful. The urge toward a higher scale of current expenditure is very great. Socialists point to corporate surpluses as an example of the manner in which capital will be saved under socialism, but it must be remembered that these corporate savings have been accumulated under the stimulus of private profit. It is true that the Russian dictatorship did succeed marvelously well, under the first Five-Year Plan, in accumulating capital at the expense of the standard of living. Such a policy would be far more difficult in a democratic socialist society.

2. Will a democratic socialist society, with its tendency toward equality in distribution, be able to motivate men of ability as effectively as the existing order? The existing order offers large rewards for superior ability. To be sure, it would be very far from the truth to say that rewards are accurately proportionate to ability in the existing order. Moreover, it must be admitted that although under existing conditions able men demand high earnings, it does not necessarily follow that they would withhold their services if lower rewards were offered in a society where there were no better alternative opportunities. Leadership is to a great extent its own reward. Private enterprise, however, demands resourcefulness, initiative, and close attention to detail to a degree not likely to be equaled by a socialist order.

3. Will a democratic socialist society be able to get as productive results from its wage-earners generally as does the existing order? The history of self-governing workshops—a very extensive one both in England and in the United States—is far from favorable on this point. Nor is the experience of municipal or national governments with their employees encouraging. The post-office service (if we except the numerous small sinecure offices) appears, however, to have attained a high degree of efficiency in the United States; but the postal employees do not control the administration, whereas under socialism the whole body of public employees would control the government. Would it be likely that such a government could impose hard and rapid work on its constituency?

4. Will a democratic socialist society be able to secure as effective administration of industry as private enterprise? It is clear that there will necessarily arise a wide divergence of opinion among the members of a socialist society as to what might be done and how. There will always be

a powerful minority striving to discredit the majority group and to oust them from control. Now a private corporation has this advantage: it can always carry out its policy in a single-minded and consistent fashion. If the minority stockholders do not like it, they have only the recourse of selling out and starting a new business of their own. A democratic socialist society would always be hampered in its administration of industry by the squabbling of conflicting groups.

5. Will a democratic socialist society be able to place all grades of workers in the various occupations as effectively from the standpoint of finding the right job for the right man, and at the same time leave each individual as much freedom of choice in the matter of occupations as the present order? A socialist society will certainly have majorities and minorities. Will not the majority seek to take all the more desirable places regardless of merit? To be sure, in the present order rich men's sons are often given desirable positions for which they are unfit; yet a firm that goes very far in that direction is likely to find itself beaten by a more vigorous and efficient competitor.

6. Will a democratic socialist society be able to allocate the productive forces of society in the different channels of production more effectively than the present order? In other words, will it be able to substitute a better guide for production than market price, or might market price itself prove a safer guide for production under a socialist society, with more equal distribution of incomes? Market price is certainly not an *ideal* guide for production. The demands, not only of the rich but of all classes above the poverty line, turn a considerable amount of our energy toward the production of goods and services which many would regard as almost completely lacking in social utility. There can be no doubt that market price can be an ideal guide for production only in a society which *utilizes* its incomes in a wise manner. But a society with an equal distribution of income is likely also to utilize its income foolishly and wastefully. This is finally, then, a question of education. A wiser expenditure of income by all classes in society is sadly needed.

WASTE IN THE PRESENT ORDER

On the other hand, we cannot be blind to the fact that the present order is very far from perfect. The gravest charge which can be directed against it is its failure to provide reasonably stable employment. Instability of employment is due, in part, to curable defects in the structure (for example, the banking system): in part to the inadequate develop-

ment thus far of compensatory controls designed to stabilize the flow of purchasing power (for example, central bank policy and fiscal policy); in part to the selfish policies of special-interest groups, which tend to make the system unworkable; and in part to the very nature of a system of free enterprise operating on a price basis. As we have seen, however, both booms and depressions—despite the waste and disorganization to which they give rise—have been important elements in economic progress. That they should be moderated is scarcely open to debate. It is very doubtful that they can be wholly eliminated in a capitalist society and equally doubtful that complete elimination would be a social gain. The unfortunate effects of these fluctuations can and should be ameliorated by social-security measures. In this manner a workable compromise may be secured between progress and security.

VALUE AND DISTRIBUTION IN A SOCIALIST ORDER

One other question remains to be considered. In a socialist society would there be any need to formulate principles relating to value, price, and distribution? Would such a society be one in which the products would be allocated by some central authority, or would it produce commodities for exchange at a value or price? Thoroughgoing state communism would presuppose the former. But if under socialism there is to be such a system as exchange and price, it is clear that the principles of price and value will be fully as important as in the present order. A socialist society would probably seek to establish a price corresponding to the cost of production. But to ascertain this cost of production would be as intricate a subject under a socialist order as under the present order. Certain socialists who adhere to the labor theory of value contend that the cost can be measured by the hours of labor required to produce the commodity. But it takes widely different numbers of hours to produce a bushel of wheat in Kansas, in Montana, and in New York. Economic rent thus emerges under a socialist order as truly as under the present order.

Nor is the matter in other respects as simple as the labor-value theorists indicate. An oak tree one hundred years old and a bushel of tomatoes may both be produced by the same number of hours of labor; but in the case of the oak tree the product does not emerge for one hundred years, whereas in the case of the tomatoes it emerges in a few weeks. One involves putting in labor for which there is no return for a hundred years. To put labor to work at something for which you can

get no return for a hundred years is certainly a cost to a socialist society as well as to private individuals, a cost of waiting in addition to the cost of labor involved. The same argument, of course, will apply to all products made by elaborate machinery and plant. A socialist government could not place on a parity two products of which one requires a capital outlay of \$5,000,000 and the other an outlay of \$5000, even though the same labor (including the pro rata share of the labor cost of constructing the capital equipment) were required in both cases. Capital is scarce, and if one industry gets more, another gets less. Thus interest emerges in a socialist society, and the cost accountants will have to consider it in determining the prices of different goods. The conclusion is inevitable that principles of value applicable to the present order would apply also in large part to a socialist order.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The English Wholesale Co-operative Society owns and operates a considerable number of manufacturing establishments. Its main business, however, is to supply goods at wholesale for the local consumers' co-operative stores. Suppose all manufacturing were brought under the control of the co-operative movement. Would interest be eliminated? profits? rent? Explain. Would a society with a universally developed co-operative system be more nearly like capitalism or like socialism? What points of similarity would it have to each? Would such a society have as great difficulty with respect to capital accumulation, efficient management, and discipline of labor as socialism would have? Explain.

2. "Higher real wages for the skilled labor, together with the increasing accumulation of property by wage-earners, will be likely to prevent the socialist parties from attaining a majority." Discuss.

3. "Should socialists actually gain majority control of Parliament in England for any considerable period, they are likely to find socialization of property ineffective. They are not likely to go far in this direction, but will then fall back upon the taxing power of the state. The net effect will be a socialization of income instead of a socialization of property." Do you agree? Why, or why not?

4. After the war the railroad unions sponsored the Plumb Plan, which provided for the nationalization of the railroads and the leasing of them to a corporation managed by fifteen directors, five appointed by the President of the United States, five elected by the railroad managers, and five elected by the railroad workers. Which form of socialism does this plan resemble?

5. Distinguish between the philosophy of the single tax and the philosophy of socialism.

CHAPTER XXXII · Some Economic Aspects of Taxation



GOVERNMENT FINANCE AND ECONOMIC PRINCIPLES

The study of the financial operations of government is called *public finance*. From the administrative point of view we may regard any governmental organization as a business undertaking. As such it bears many resemblances to a private business. Systems of accounting and control must be established, labor and commodities must be purchased and applied to the production of other goods and services, income and expenditure must be balanced, and the personnel of the working force must be organized and directed. Moreover, many of the important economic principles (the principle of diminishing returns, for example) that must be observed by the private business are significant for the conduct of governmental business also. On the other hand, there are important differences between private business and governmental business. Private business seeks the highest total profit as its goal; the government seeks the maximum public welfare. The income of private business is derived from the sale of commodities or services; the income of a governmental organization is derived chiefly from taxes imposed on private wealth and incomes. It is a relatively easy matter for the manager of private business to determine whether it is profitable to produce a given commodity or service, but it is not always possible to discover whether a given governmental service is worth what it costs the taxpayers.

From the social point of view governmental business, or, more briefly, "the government," may be looked on as a means for accomplishing certain ends that we have decided should be taken from the hands of private business, or for the provision of certain services that no private business would find it profitable to carry on. The rule of private business is that only those commodities and services will be produced for which the consumers are willing and able to pay a price at least sufficient to cover the costs of production. But certain necessary public services, such as military protection and the maintenance of law and order, cannot be provided by a private concern, or at least they will

not be provided in a satisfactory manner. Another rule of private business is that only those who are able and willing to pay the necessary price will be provided with the commodity or service. But there are many services, such as elementary education, sanitation, police protection, and the public highways, which we believe should be available to everyone regardless of ability to pay; hence we turn these services over to governments, in whose hands they are provided for the use of everyone who may properly be said to require them, and their cost is met by taxation. There are exceptions to this rule (as we shall see later) when the users of the certain governmental services are required to pay all or a part of the costs of providing them.

When the government renders a service that is offered without charge to all who may require it and then raises the funds to defray its cost by a levy on the wealth and income of the citizens, it usually happens that the distribution of the wealth and income of the people is affected to a greater or less extent. Private incomes are very unequally distributed, and the same thing is true of private wealth. Many persons who require large amounts of governmental service are unable to pay proportionate amounts of taxes; they receive proportionately more from the government than they contribute in taxes. Others, who own large amounts of wealth or receive large incomes, contribute proportionately more to the cost of maintaining the government than they receive from it in services. It must be noted that the difference between contribution and benefits received is relative. The receiver of a large income could not do without governmental service; his income would vanish were it not for the activities of the state; in a certain sense it therefore is correct to say that his entire income depends upon such governmental services as the maintenance of laws controlling the use of property. He does not contribute absolutely as much to the support of government as it is worth to him; but he contributes in most cases much more than the receiver of a small income, whose well-being likewise is completely dependent upon state activities. Since the effect of governmental services under our present system is to cause some to contribute relatively more than the relative benefits received by them, and others to contribute less, we see that taxation has an important effect on the distribution of wealth and income.

THE NATURE OF A TAX

A tax is defined as a compulsory payment collected by the government from individuals and corporations, without reference to benefit, for the support of governmental operations.

In the first place, a tax is a compulsory payment. Sometimes the government offers goods for sale in the market like any other seller. Those who wish may avail themselves of these goods by paying the price; those who do not wish to do so may escape payment. Such governmental charges are *prices* and not taxes.

In the second place, a tax is always collected for the support of governmental functions, the courts having consistently held that the power of taxation may not be employed to provide subsidies that merely enrich some individuals at the expense of the general body of taxpayers. The objects upon which tax money is expended must always be public purposes. They must always be objects that contribute to the general welfare of the community. Under ordinary circumstances it would not be legal to levy a tax for the purpose of building a house for a particular person, since the service derivable from the house would benefit the person to whom it was given without benefiting the community at large. Public services, such as highways, education, and police, are paid for by taxation because they confer a general benefit on all the people. The gain of individuals from these services is either incidental to the public benefit or it is received by individuals as members of a community in common with all other persons in the community.

In the third place, a tax that has been legally imposed must be paid by the individual, regardless of the benefit he may receive from its expenditure. A wealthy taxpayer cannot refuse to support the public schools because he has no children to send to them; nor can he get his taxes reduced on the ground that the benefit which the school system confers on him or his property is not proportionate to the amount of his taxes.

NONTAX REVENUE

A certain amount of revenue is raised by governments from sources other than taxation. These nontax sources may be classified as follows: (1) prices (that is, receipts from commercial enterprises), (2) fees, (3) special assessments, and (4) fines and penalties.

The term "prices" applies to revenues received by the government from the sale of commodities and services. The Federal government

sells lumber from the national forests and power from its hydroelectric plants; city governments often own and operate water systems and charge the users according to the amount of water consumed. Sometimes these prices are essentially the same as the prices charged by private businesses. Lumber from national forests is sold in the competitive market for whatever price it will bring. In other cases the government monopolizes the service and charges prices that may be either above or below the prices that would prevail if there were competition. Like other monopolists, the government may discriminate between different users.

A fee is a payment imposed on individuals by the government in respect to some service performed by it for the benefit of those individuals. It differs from a price in that the service for which it is charged is more important for the public at large than is the service for which a price is charged. The service is usually undertaken by the government for the purpose of regulation and control. Typical local fees are those charged for the inspection of building plans, boilers, and electrical wiring, and for the recording of deeds to real property. The amount of the fee should not exceed the cost of rendering the service, and it may be less than the cost; in practice, however, it sometimes exceeds the cost of the service.

A special assessment is a kind of fee. It is charged when a city (it may be imposed by other governmental organizations) improves the streets or parks. It is paid by the owners of the property benefited by the improvement. Here again a public purpose is involved, and therefore it is proper to compel the owners of all taxable property to contribute. It would be possible to finance street improvements out of money raised by taxation, as rural highways are financed; but since it is generally true that such improvements confer pecuniary benefits on the owners of the adjacent real estate, a special charge is properly collected from them. In certain instances, of course, when the owner of the real estate is benefited little or not at all, all or a part of the cost of the improvement may be borne by the taxpayers at large.

KINDS OF TAXES

There are many different kinds of taxes. Everyone is familiar with the local taxes on real estate and personal property and with the Federal income tax. All these are *direct* taxes. On the other hand, there are many taxes, such as customs duties and the duties on tobacco and

other commodities, that are *indirect* taxes. Direct taxes are borne by the receivers of the incomes and the owners of the property upon which they are levied; indirect taxes are added to the prices of the commodities on which they are imposed and are borne by the consumers of these commodities.

Another classification of taxes distinguishes them according to the base upon which they are levied. The base of a tax is the taxable thing according to the quantity of which the amount of the tax is determined. For example, the base of the local tax on real property is commonly the sale value of the taxable property. If the sale value of the property is large, the tax will be large; if it is small, the tax also will be small. In the United States, taxes are principally imposed according to the following bases: (1) the value of real estate, (2) the value of personal property, (3) net incomes, (4) gross earnings, (5) commodities, (6) privileges, and (7) acts. All these bases are self-explanatory except possibly the last. Taxes on acts comprise the duties payable on certain legal transactions (such as the recording of wills and deeds), on powers of attorney, on the making of contracts, and the like. As sources of revenue they are unimportant, and the general opposition to them by the public indicates that they will remain so.

THE INCIDENCE OF TAXES

When a tax is levied upon any kind of property or income, it must be paid by the owner of the thing taxed. But the owner may be able to recover the tax from someone else who is a buyer or a seller of the taxable thing. When a person on whom rests the legal obligation to pay a tax passes it on to someone else, he is said to *shift* the tax. The person who finally pays, or the property from the income of which the tax is ultimately deducted, is said to *bear* the tax, and the *incidence* of the tax is said to be either on this person or on the property in respect to which the tax must be paid.

When a tax is imposed on producers' capital or on a particular commodity, it may have effects similar to those that follow from an added cost. For example, if a tax is imposed on a commodity and made payable by the producer, it may be passed on to the consumer, just as a rise in wages may be. But the notion that all taxes on business or on incomes from business are shifted to the consumer is erroneous. The first condition necessary for the shifting of a tax is that there shall be some untaxed field to which the capital, labor, and land subjected to

the tax can withdraw. If the producers of cloth were taxed on each unit turned out, the only method by which they could pass the tax to the consumer would be by raising the price. This would cause a decline in sales, and unless the producers were prepared to accept a reduction of output they would not be able to shift the tax. A reduction in output can come in two ways: the factors may be withdrawn to some field where taxes are lighter, or the factories may run at less than capacity. If the former method is open to the producers, they will be able to shift the entire burden of the tax at once; if it is not, they may gain an advantage by curtailing production, raising the price, and waiting for the natural growth of demand to absorb the entire output of their factories again. The two methods are evidently the same, for the latter involves for its success the exclusion of new capital and the diversion of it to untaxed fields. Since taxes on commodities either do not hit many sorts of goods, or constitute unequal burdens on different commodities, most of them can be shifted after a longer or shorter period of time.

It is said that a tax on economic rent cannot be shifted to the consumer of the products of the land that yields the rent; but if the tax is imposed on the rent of a tract of land only when it is used to grow wheat, and if at the same time the owners can derive approximately as great a return by growing some other crop, the tax will be shifted. The tax reduces the rent the owners receive, and if this reduction is so great that they can obtain larger returns from their land by producing some other crop, they will produce that crop and not wheat. The supply of wheat having fallen off, its price will rise, and the consumer will bear the tax. But if rents from all the uses of the land are taxable, the tax does not affect the owner's choice of products; the relative supplies will remain as they were before the tax was imposed, and the tax will not shift.

In the second place, a tax that hits only surpluses cannot be shifted. By a surplus we mean a gain in excess of the expenses of production, including imputed expenses for the entrepreneur's land, capital, labor, and management. Fortuitous profits above the marginal cost of production, good and bad years taken together, may be taxed away without affecting the supply of the commodity. This statement does not mean, of course, that the state can take all gains from investments in excess of the riskless rate of interest; a part of these surpluses must be used to offset unavoidable losses.

A tax on net income usually cannot be shifted if it hits all occupa-

tions or industries alike. Since the tax does not increase or decrease with the amount of goods produced or with the amount of work done by the taxpayer, but with the net receipts from such economic activity, the individual can escape it only by causing his net income to decline. This no one would do as long as he had in mind the maximization of his income; for a hundred dollars received is better than fifty dollars, even though the government takes 10 or 20 per cent of either. The foregoing reasoning applies to corporations as well as to individuals. It is possible that an income tax at extremely high rates would deprive the taxpayers of so large a part of their incomes from labor and capital that industriousness and saving would be discouraged, and that, as a result, the social income would be reduced and the distribution of that income changed. But these effects would not prove that the tax had been shifted, because the tax would still be a deduction from the incomes of those on whom it was imposed.

When a tax hits margins it is more likely to shift than when it hits nonmarginal incomes, firms, etc. A business tax that impinges on marginal producers in any field will drive them out of that field if there is any other to which they can go and escape the tax. But even though the tax is not general, it is less likely to be passed on to consumers if it is imposed on firms that are making a pure profit. These firms receive for their goods a price that more than covers all expenses (including imputed expenses for proprietors' capital), and they will be less likely to desert the taxed trade when compelled to pay a tax than would the marginal firm.

A tax on monopoly revenue cannot be shifted. In other words, if the tax subtracts a certain percentage from the excess, above cost of production under competitive conditions, which a monopolist receives by virtue of exclusive control of supply, the monopolist cannot compel consumers to reimburse him. By assumption the monopolist is already receiving the price that will yield the maximum profit; and if he raises the price, his maximum net income will be reduced and he will still have to pay the tax at a certain percentage rate on the monopoly revenue. On the other hand, if the tax is imposed on the commodity produced by the monopolist and if he is required to pay it, the tax will probably shift in part, because the monopolist, like any other producer, can withdraw his capital and reduce the output. He thus can escape part of the tax and at the same time raise the price to the consumer. The capital withdrawn can be invested at the competitive rate free from this special commodity tax.

The details of the process of shifting are very complex. Here we have attempted to set forth only the most general principles. A tax is not easily shifted (1) when it is universal—when it hits all people and all business alike; (2) when it hits net income; (3) when it hits surpluses and does not impinge on margins.

In the preceding discussion we have assumed that the agents of production are completely mobile. This assumption must now be qualified. If a commodity tax is imposed on producers, they may bear it for a time because they are unable to change the supplies of goods they are bringing to market. Their capital may be highly specialized, or their land not easily turned to other uses. But all capital wears out in the long run, of course, and when the wearing-out process has begun to reduce the supply of the commodity, the tax will begin to shift. Even the labor of the proprietors may not be readily adaptable to other types of production. The supply of the product being fixed, there is no method by which the producers can force the tax on the consumer.

The speed and ease with which a tax on commodities can be shifted to the consumer is dependent upon the elasticity of demand for the product. If the demand slopes very gradually, it may require a large reduction of supply to raise the price by the amount of the tax; if it is very steep in slope, only a small reduction will be necessary.

A tax may be shifted backward as well as forward; for example, if a heavy commodity tax is imposed on the producers of a given good, they will try to pass it to the consumers by reducing output. But a reduction of output not only curtails the amount of the finished commodity; it also curtails the demand for labor and raw materials. If the labor used in making the commodity can easily move to other occupations, and if the land and capital with which the raw materials are produced can easily be turned to other uses, then the consumer will have to pay the tax; but if the labor and the land are not industrially mobile, then either or both of them may be compelled to submit to lower returns and thus be compelled to bear part of the burden of the tax.

Finally, we must distinguish between the shifting of a tax and its *effects* on property values and individual wealth. If a very heavy tax were imposed on grocers it would be shifted, under ordinary conditions, because some of the people engaged in the grocery business would no longer be able to meet their costs and therefore would be forced to withdraw. In the process of withdrawal many would lose money. They would be forced to sell their fixtures at a sacrifice, to

abandon such patronage as they had built up, and to seek different occupations at lower wages. But the tax might still be shifted in its entirety to the consumer. Those who remained would be able to cover their costs, including the tax, and the prices of goods would rise enough to cover the tax as well as other expenses. The loss sustained by the owners of abandoned grocery stores would be an effect of the tax; its incidence would still be on the users of the goods.

THE CAPITALIZATION OF TAXES

When a tax hits a durable good, either in its primary incidence or ultimately, it tends to be capitalized. This phenomenon can be illustrated best in the case of land. Let us assume that a parcel of land yields \$100 rent annually and that the rate of interest is 5 per cent. The land will then be worth \$2000. Now let a tax of \$20 a year be imposed on the rent. The rent which any owner can count as net will be reduced to \$80 and the value of the land to \$1600. To put the same calculation in another way: the capital value of the land before the tax was levied was \$2000; the capital value of the tax (\$20 a year at 5 per cent) is \$400; this amount is subtracted from the original sale value of the land, leaving a new capital value of \$1600.

The first condition necessary for capitalization has already been stated: the thing taxed must be durable. If it is not, and if supplies are constantly coming on the market, these supplies will be checked when the tax is imposed, and the burden of the tax will be passed forward to the consumer; if the durability of the good is considerable but not comparable to land or to land improvements, then there may be a partial capitalization of the tax. The second condition is inequality. It is often said that taxes on rent are always capitalized, but this conclusion is hardly justified by theoretical reasoning. The problem turns on the question whether the first purchaser of the land after the tax has been imposed will be able to compel the seller to accept a reduction in its price equal to the capitalized value of the tax. The investor in land, like any other buyer of capital goods, seeks to put his money where, other things being equal, it will earn the maximum return. If he finds that land has been subjected to a new tax and that other permanent producers' goods have not, he will be able to force the original landowner to accept the loss of \$400 by threatening to buy other property not affected by the tax. Concretely, he will say to the landowner that he and other investors can invest \$2000 in machinery, in buildings, or

in merchandise and receive \$100 annually free from any tax deduction, and the landowner will be compelled to accept \$1600 if he wishes to sell. On the other hand, if the prospective purchaser finds that all alternative investments are affected by the tax to the same extent as land is, he will be unable to compel the landowner to accept \$1600, and the tax will not be capitalized.

THE AMOUNT OF TAXATION

The amount of taxation imposed by any government is determined in large part by the amount of its expenditures, but not all the money receipts of any government in a given year are derived from taxes. Part ordinarily comes from fees, part from public enterprises such as waterworks and the post office, and part from loans. That the first two constitute independent sources of revenue is easily understood. Loans, on the other hand, appear to be mere anticipations of revenue from taxation or from nontax receipts. This is usually the case when borrowing takes place for the purpose of financing permanent improvements, such as public buildings, roads, sewers, and waterworks; the government borrows to obtain immediate funds, and the bonds it issues are retired later by payments from the charges made to consumers or by funds raised by taxation. When loans are made to meet the expenditures arising from emergencies, such as wars, governments frequently attempt to wipe out at least part of the indebtedness by taxation when the emergency is past. But they do not always do so; hence borrowing for these purposes may cause an increase in receipts which is not offset later by taxation. Moreover, although the original indebtedness for permanent improvements may be wiped out, the time comes when the improvement must be replaced and further loans are made for that purpose. Public debts thus may become permanent. When, however, the public debt becomes unmanageably large, repudiation may be resorted to, or by inflation the government may escape all or a part of the burden of the debt. Yet in spite of these discrepancies the amount of public expenditures is a good measure of the amount of taxation required in any community.

The control of the amount of taxation depends primarily, then, upon the control of the amount of public expenditures; but the determination of the legislature to spend is partly controlled by the amount of taxation that it is thought advisable to impose. Hence we conclude that the volume of public expenditures and the volume of taxation mutually

govern each other; if it is easy to raise large amounts by taxation, public expenditures are likely to be large; if there is great need for public expenditures, the amount of taxation is likely to be great. In the first place, the ability of the people to sacrifice a portion of their incomes for the support of public services must be considered (a population whose income is so small that it barely affords the necessities of life cannot contribute large amounts to the support of governmental services). In the second place, the utility that can be obtained from the expenditure of public revenue also is a determining condition. Consider the case of public highways. Before the automobile became common, graveled or dirt roads were generally thought to be adequate for rural districts; but with the growth in the use of motor cars the demand for highways with a continuously smooth and firm surface became insistent. Money was borrowed and taxes were imposed to raise the funds for improving them. In this case the additional utility derivable from better highways dictated the increase in taxation.

But the need for improvements alone does not control. Some communities are too poor to support good roads, good schools, and expensive public buildings; hence, notwithstanding the utility of these goods and services, they are not provided, and taxation remains relatively low in absolute amount.

A third factor that enters into the determination of the amount of taxation is the efficiency of government. If the administrative departments of government are efficient, people are more likely to consent to the imposition of heavy taxes than if they are corrupt and inefficient; yet we find that taxes are sometimes very high in states and cities where the standards of efficiency and honesty among public officials are very low. This is because there are still other forces at work to determine the amount of taxes. Fourth, heavier taxes can probably be imposed if the tax system includes many indirect taxes than if it is composed chiefly of direct taxes. The reason is obvious: indirect taxes are shifted to the users or consumers of the articles on which the taxes are first imposed; they are not paid ultimately by those who must first advance them. The ultimate payers often are not cognizant of the fact that they are paying a tax when they buy goods or ride on street cars, and when they are aware of it they may not know the exact amount of their payments. A user of tobacco ordinarily will pay more willingly a tax of \$10 when it is added to the price of his cigars or cigarettes than he would if it were assessed against his income or his real estate.

The attempt has often been made to lay down rules for determining the proper amount of taxation to impose in any given community. As we have seen, the amount depends primarily upon the available income from which taxes can be paid and upon the value of the services the public receives from public expenditures. This much is generally admitted. But the foregoing statement is indefinite. It does not tell us whether it is good economy to take 5 per cent or 20 per cent of the income of the people for the support of governmental services. The impossibility of laying down definite statements is inherent in the nature of the problem. The conditions that have been stated in the earlier portions of this section are not constant. The valuation that taxpayers set upon public services varies with the changing conditions of industrial life, with changes in the opinions of the voters as to the proper degree of state interference, and with changes in the efficiency and technique of public services. The income available for taxation also is variable, changing not only with the variations in the total income but also with the variations in its distribution.

INCREASE IN TAXATION

Any change in the amount of public expenditures is certain to affect the amount of taxation. And for more than a century the amount of public expenditures has been persistently on the increase. In the United States this increase has been relatively slow in times of peace; but two great wars, the Civil War and the World War, have been the occasions, if not the causes, of sharp upward trends in taxation. That a great war should cause an increase in the expenditures and taxation of the national government is easily understood, but that it should also be followed by a rise in local taxation is not easily explained; nevertheless the evidence is clear.

Before the World War the expenditures of the Federal government were steadily but slowly increasing, and the same was true of state and local expenses. In part this increase was due to increasing population, but the growth in governmental functions also was an important factor. The per capita receipts from taxation by all the governments, which made up the bulk of all expenditures, rose from \$17.30 in 1890 to \$22.70 in 1913 in terms of dollars of constant purchasing power. By 1923, however, these taxes, similarly adjusted, amounted to \$42.20.¹

¹*Cost of Government in the United States, 1925-1926*, pp. 70, 76. National Industrial Conference Board.

TABLE 50. Gross Governmental Expenditures of the United States, Fiscal Years 1923 to 1934¹

YEAR	TOTAL	FEDERAL	STATE	LOCAL
Amounts in Millions of Dollars				
1923	9,920	3,885	1,242	4,793
1924	10,974	4,121	1,432	5,421
1925	11,126	3,765	1,532	5,829
1926	11,616	3,936	1,539	6,141
1927	12,179	4,069	1,656	6,454
1928	12,609	3,970	1,826	6,813
1929	13,048	3,932	1,990	7,126
1930	13,428	4,141	2,223	7,064
1931	13,516	4,172	2,367	6,977
1932	14,453	5,225	2,322	6,906
1933	13,316	5,264	2,141	5,911
1934	15,496	7,207	2,132	6,157
Dollars per Capita				
1923	88.94	34.83	11.14	42.97
1924	96.94	36.40	12.65	47.89
1925	96.86	32.78	13.34	50.74
1926	99.68	33.77	13.21	52.70
1927	103.04	34.43	14.01	54.60
1928	105.20	33.12	15.24	56.84
1929	107.37	32.36	16.38	58.64
1930	109.09	33.64	18.06	57.39
1931	108.90	33.61	19.07	56.21
1932	115.65	41.81	18.58	55.26
1933	105.88	41.85	17.02	47.00
1934	122.38	56.92	16.84	48.62

From 1923 to 1929 the gross expenditures of the Federal and the state governments remained relatively constant (see Table 50), while local expenditures increased over 40 per cent. From 1929 to 1934 the outlays of all three divisions—Federal, state, and local—increased, but not by any remarkable amount. Since 1933, however, Federal expenditures have increased greatly (Table 51). The rise from 4.8 billions in 1933 to almost 8.0 billions in 1937 has been brought about by several causes. General payments were affected by such expenditures as the adjusted service compensation of World War soldiers in 1936–1937 as well as increases in the regular departments. However, expenditures for recovery and relief have been the principal causes of added outlays.

¹ *Cost of Government in the United States, 1933–1935*, p. 7. National Industrial Conference Board, 1936.

TABLE 51. Receipts and Expenditures of the National Government of the United States, 1930-1937
(in Millions of Dollars)

YEAR ¹	RECEIPTS ²					EXPENDITURES ³							
	Total	Income Taxes	Miscellaneous Internal Revenue	Customs	Agricultural Adjustment Taxes	Nontax Receipts	Total	General	Recovery and Relief	All Other			
							Total	Interest	Other	Total	Relief	Public Works	
1930	4177	2411	628	587		551	3994	659	3335				
1931	3190	1860	569	378		382	4092	612	3480				
1932	2006	1057	504	327		117	4862 ⁴	599	3374	768			
1933	2080	746	858	251		225	4845 ⁴	689	2715	1277			
1934	3116	818	1470	313	353	162	6745	757	1984	4004	1844	645	1515
1935	3800	1099	1657	343	521	180	6803	821	2325	3657	2342	1020	294
1936	4116	1412	2024	387	77	216	8476	5195	749	4436	3291	985	213
1937	5225 ⁵	2158	2181	486		141	7988	866	4275	2846			

¹ Year ending June 30.

² From *Daily Statement of the United States Treasury*.

³ From *Daily Statement of the United States Treasury*, for the years 1930-1931 and 1937; for the years 1932 and 1933, from *Federal Reserve Bulletin*, March, 1934; for the years 1934-1936, from *Federal Reserve Bulletin*, February, 1937.

⁴ Total contains trust-fund and contributed-fund expenditures not shown separately.

⁵ Total contains, in addition to items shown separately, 252 million dollars Social Security taxes and over 6 million dollars miscellaneous taxes.

The increase in the amount of taxation does not measure the burden of taxation. During the past four decades national income has been increasing and a larger amount has been available for paying taxes. The burden of taxation depends upon the proportion of the income taken by taxes and not on the absolute amount of taxes. In Table 52 the amounts of taxation raised by the three governmental divisions are compared with the national income.

TABLE 52. Taxes and the National Income, 1890-1934¹

YEAR	NATIONAL INCOME (IN BILLIONS OF DOLLARS) ²	PERCENTAGES THAT TAXES ARE OF THE NATIONAL INCOME			
		Total	Federal	State	Local
1890 . . .	12.1	7.2	3.1	0.8	3.3
1903 . . .	20.5	6.7	2.5	0.8	3.4
1913 . . .	34.4	6.4	1.9	0.9	3.6
1919 . . .	68.3	10.9	6.6	0.8	3.5
1922 . . .	57.3	13.1	6.1	1.5	5.5
1923 . . .	65.3	11.1	4.6	1.4	5.0
1924 . . .	67.8	11.5	4.7	1.5	5.3
1925 . . .	72.3	10.9	4.1	1.5	5.3
1926 . . .	74.2	11.6	4.3	1.7	5.6
1927 . . .	75.2	12.0	4.4	1.8	5.8
1928 . . .	76.5	12.2	4.2	2.0	6.1
1929 . . .	78.6	12.4	4.2	2.1	6.1
1930 . . .	72.9	14.1	4.8	2.4	6.9
1931 . . .	61.7	15.1	4.4	2.9	7.8
1932 . . .	48.1	16.8	3.7	3.4	9.7
1933 . . .	44.9	16.7	4.0	3.4	9.4
1934 . . .	50.2	17.5	5.8	3.5	8.3

It is shown by Table 52 that from 1890 to 1913 the percentage of the national income taken by the taxation of the national government declined, that state and local taxation increased slightly, and that the total taken by all taxation declined. The World War was the cause of an increase in the burden of national taxation, as is shown by the rising percentages of the national income taken by Federal taxation after 1913. But the interest charges, debt repayments, and pensions from 1922 to 1929 were not the sole reasons for the increase. The ratio of state and local expenditures to national income also increased.

¹From *Cost of Government in the United States, 1925-1926*, p. 77, for the years 1890 to 1919 inclusive, and from *Cost of Government in the United States, 1933-1935*, p. 30, for succeeding years. National Industrial Conference Board, 1936.

²Income paid out from 1929 to 1934.

It must be emphasized that the percentages of the national income taken by taxation as shown in Table 52 depend both for their magnitude and for their trend upon the amount of the national income. There is no census of income and no other source from which the total can be accurately ascertained. The figures given are estimates, and, as is usually true of estimates, authorities are not entirely in agreement. The figures given for income are conservative by comparison with the computations of other competent statisticians; yet even on the basis of these conservative figures the burden of taxation in this country is evidently not crushing. Moreover, it must be remembered that the ability to spare private income for the support of public services increases more rapidly than the national income.

Nevertheless the fact remains that governmental taxation and expenditure are increasing. The causes are not the same for all three forms of government. The principal object of Federal expenditure down to 1933 was payment for past wars and preparation for future wars; for the state and local governments the principal objects requiring taxation were education and highways, followed in order of importance by protection to persons and property and by social relief.

After the return of the Federal government to a peace-time basis, from 40 to 50 per cent of its expenditures were for repayment of, and interest on, the national debt. Since the debt is held at home, these disbursements are not a "burden" in the same sense that expenditures for general government and the support of the army and navy are burdens. The money to make payments on debt is collected by taxation from the citizens of the United States, who also hold the bonds. The entire transaction is a transfer of wealth from the taxpayers to the bondholders, which constitutes no present burden upon the total income of the nation.

The increase in public expenditures is a phenomenon that has manifested itself for more than a century. It is not confined to certain countries. In different countries somewhat different objects of expenditure are responsible, though the increase is generally conceded to be due to (1) the increasing cost of armaments, and war debts, (2) technological progress, and (3) the growth of democracy. The progress in technology makes possible services that could not have been provided at an earlier time; the progress of democracy gives rise to a demand for provision at public expense of better services and of new services that progress in technology makes possible. In a democracy the mass of the voters, who on the average own very little property and receive rela-

TABLE 53. Net Public Debt in the United States, 1913-1936¹

YEAR	TOTAL (MILLIONS OF DOLLARS)	UNITED STATES GOVERNMENT (MILLIONS OF DOLLARS)	STATES, COUNTIES, AND CITIES, ETC. (MILLIONS OF DOLLARS)
1913	4,874	966	3,908
1914	5,247	967	4,280
1915	5,651	969	4,682
1916	6,052	970	5,082
1917	8,141	2,711	5,430
1918	17,627	11,933	5,694
1919	31,076	25,086	5,990
1920	30,520	23,854	6,666
1921	30,775	23,389	7,386
1922	30,854	22,289	8,565
1923	30,688	21,596	9,092
1924	30,602	20,596	10,006
1925	30,772	19,695	11,077
1926	30,553	18,758	11,775
1927	30,238	17,523	12,715
1928	30,057	16,497	13,560
1929	30,234	15,755	14,479
1930	30,516	14,950	15,566
1931	33,065	16,084	16,981
1932	35,826	18,592	17,234
1933	38,820	21,507	17,313
1934	42,430	25,226	17,204
1935	43,206	26,060	17,146
1936	47,762	30,880	16,882

tively low incomes, have the power to vote expenditures and levy taxes. And since direct taxes bear most heavily on the receivers of large incomes and the owners of property, there is constant pressure for the provision of additional services at public expense.

GOVERNMENT DEBT

Governments borrow when public revenues are inadequate to finance the expenditures believed by legislatures necessary to carry on the functions of government. There are three principal occasions that usually give rise to this inadequacy: (1) war, (2) permanent public works, and (3) catastrophic events, such as severe depressions, floods,

¹From the Annual Report of the Secretary of the Treasury, June 30, 1936, p. 465. The data are for tax-exempt securities. They do not include debt (1) held by United States trust funds, (2) owned by the United States or by governmental agencies, and (3) held in sinking funds. The debts of territories and insular possessions are not included.

etc. The last give rise to the need of public support of large numbers of the people and at the same time cause a severe falling off of revenues. Sometimes, of course, a government may borrow to meet the ordinary expenses of government under conditions of prosperity. This procedure is not regarded, however, as sound financial practice.

The need for borrowing in time of war is self-evident. From Table 53 it can be seen that in a single year, 1918-1919, the debt of the Federal government advanced from about 12 billion dollars to 25 billions. In this same year the receipts of the national government from taxation amounted to over 4.5 billion dollars. It is conceivable that a much larger amount could have been raised by taxation, but it is highly improbable that the amount could have been increased by 13 billions.

Borrowing to construct public works often stands on a different footing. Canals and power plants are very often regarded as improvements that will earn enough revenue to carry the interest charge on the loans and possibly retire the principal. Public buildings are built and parks are acquired for the use of the public in the future, and it is thought proper that those who have the benefit should carry the interest and amortization charges.

Never before 1930 has public borrowing to support the victims of an economic depression and to stimulate recovery been resorted to in the United States on a large scale. In 1929 the total outstanding debt of the Federal, state, county, and city governments amounted to 30.2 billion dollars. During the next seven years it increased by 18 billion, or by 60 per cent. The Federal government increased its debt by more than 15 billion dollars. During these years the revenue receipts of that government amounted to 22.5 billions, but this was wholly inadequate to provide for the large amounts dispersed through such agencies as the Federal Relief Administration and the Works Progress Administration. The receipts from many of the old taxes levied by Federal and state and local governments fell off greatly during this period and, with the scale of expenditures adopted, the only recourse was borrowing.

TESTS OF A GOOD TAX SYSTEM

There are certain qualities that a tax system should possess to make it administratively successful. These qualities or characteristics must be taken into account in any discussion of the effect of taxation on the distribution of wealth and income or on the prices of goods. In the first place, a tax system must be productive; it must yield adequate revenue

before any other test can be taken into account. Second, the system as a whole should be flexible enough to permit the fiscal authorities to adjust the revenue to changing needs. Third, the taxes that make up the system should be dependable; that is, they should be so imposed that the revenues from them can be forecast with some degree of certainty. This is desirable in order that proper adjustment of receipts to expenditures may be made in advance of the collection of the taxes. Fourth, every tax ought to be economical. The costs of assessment and collection ought not to consume a large portion of the revenue that it yields; neither ought the tax to be imposed in such a manner that the taxpayers are greatly inconvenienced or put to much expense to comply with the rules and regulations governing its payment.

In addition to these administrative tests, a tax or a system of taxes ought to meet certain nonadministrative requirements. It ought to be imposed in such a manner that the amounts of tax paid by people in different income classes conform to the sense of justice of the community in which it is levied, and it ought to conform to the political institutions of the government imposing it.

THE DISTRIBUTION OF TAXATION

Two principal theories of the distribution of the burdens of taxation have been advanced by writers on this subject. The benefit theory asserts that people should pay according to the benefits they derive from the operations of the government. An obvious, but superficial, argument in favor of this theory is that taxes are burdensome and ought therefore to be paid by those who occasion them. It is not the purpose of taxation, it is said, to favor some people at the expense of others, and to avoid this it is necessary to tax people according to the governmental expense they occasion. The first objection to the benefit theory is its lack of practicability. Much of the activity of governments is made necessary by people who are too poor to contribute their full share to taxation. One can hardly maintain, for example, that unless a person of small income can pay the cost of educating his children they should not attend the public schools, or that unless he pays his proportion of the cost of police the state should not protect him. The theory is impracticable also because it is impossible to discover the individual benefits resulting from many government services. A military establishment maintained by the national government is a necessity, but it is impossible to make even a rough estimate of the benefit received from it by any given person.

A second, and more fundamental, objection to the theory is the fact that it is based on an erroneous concept of the purpose of government. Organized society does not exist for the benefit of individuals as such, but for the welfare of the entire community. Governmental services are not designed to advance the interests of individuals except so far as their advancement is an advantage to the community.

There is, however, a certain amount of admitted justice in the theory. It is difficult to defend a tax on the people of New York to build local roads in Nebraska or a tax on the people of one city to support a fire department in another remote municipality. And since the owners of automobiles occasion the major expense of permanent highways, it is only fair that they should pay special taxes for the support of this service. On the other hand, there are many who believe that the national government should assist in the maintenance of a system of education, and it is common practice for a state to tax all the people within its jurisdiction for the partial support of the schools in communities where, without this aid, passable educational standards could not be maintained.

A second theory asserts that taxes should be paid in accordance with ability to pay. This means in its most general form that all the people living within the jurisdiction of a government should contribute to its support according to their ability as measured in terms of wealth or income. In support of this theory it is said that government is a common enterprise undertaken for the good of all. All are obliged to support it to the limit of their capacities if the need arises. In extreme cases, when public danger is great, the state may take any portion of a citizen's property needed, may circumscribe his freedom of action, and may even force him to enter the army and risk his life. Ordinarily so great a contribution is not required, but the same principle holds for the smaller need as for the greater one: taxes should be apportioned according to ability.

How ability to pay should be measured is not an easy problem to solve. Wealth is not an entirely satisfactory criterion, because many people have incomes and therefore ability to pay without possessing much wealth, and the capital values of different forms of wealth do not vary exactly with the income derived from them. Money income is deficient also as a test of ability. In the first place, some incomes are derived from the ownership of property; others are the results either of labor or of the active management of business. The former are commonly called *unearned* incomes; the latter are called *earned* incomes. To

tax two individual incomes of the same size, one of which is earned and the other unearned, at the same rate is clearly not just. In the second place, some money incomes are certain and others are uncertain; some are temporary or limited in duration and others are relatively permanent. Incomes from labor and from the management of businesses are both more uncertain and more temporary than incomes from property. Third, individuals having similar incomes may have greatly different obligations. An unmarried individual who receives \$5000 annually has greater ability to pay than another who receives the same income but who must provide for a family. These differences in ability are recognized in income-tax laws by allowing exemptions for dependents and by imposing lower rates on earned than on unearned incomes.

Still another theory, the equal sacrifice theory, which is essentially a variant of the theory of ability to pay, asserts that taxes should be so imposed as to cause each taxpayer the same sacrifice. Here the burden of taxation is measured in terms of satisfaction that must be given up through the payment of taxes. The measure of the tax burden is conceived to be subjective. One obvious obstacle to the application of this theory in practice is the difficulty in ascertaining the amount of subjective loss occasioned by a given tax payment. It is not difficult, perhaps, to convince ourselves that to take 10 per cent of the income of a mechanic who earns \$1500 a year imposes a greater sacrifice than to take the same percentage from the income of a professional man who earns an income of \$10,000 a year; but when the distinction must be made between individuals receiving large incomes the application of the principle is not so easy. Moreover, we cannot be sure that to take the same percentage of income from two individuals who have the same incomes and the same needs will impose the same sacrifice, although there could be little injustice in assuming that it would. In conclusion, it may be admitted that the theory is workable in a rough way, but a rigid application of it cannot be defended.

None of these theories is broad enough to include all the considerations that should enter into the determination of the distribution of taxation. In some cases the principle of benefit is applicable. These are chiefly instances in which the benefit is assignable to a group of people living within a city or some similar political division. Usually, too, the benefit theory is applicable when a nonindigent group receives a direct pecuniary return from a certain public service which is important to them as well as to the community at large. Hence we expect the people within a given political division to support their local government. But

when the service to be supported is of great social importance and when the benefits cannot be assigned, we must resort to the theory of ability to pay or the theory of equal sacrifice to guide us in the distribution of the tax burden.

Some taxes that have the stamp of long approval on them cannot be justified by any of the theories we have set forth. A good example is the sumptuary tax, which is levied wholly or in part to discourage the use of a luxury or of a good considered detrimental to public welfare. Taxes imposed on alcoholic beverages, which are extremely heavy as compared with other commodity taxes, are a case in point. Again, neither benefit nor ability to pay can be used to justify the heavy import duties imposed on some commodities under the policy of protection. In most communities there is an opinion, often not definitely formulated, that landowners ought to bear heavier tax burdens than the owners of other sorts of property, because the rent of land is looked upon as an unearned income. Of course, the theory of benefit enters also, because the landowner is thought to receive greater gains from the growth of the community and from public improvements than other property-owners.

In conclusion, it appears that there is no one rule that can be followed rigidly in the apportionment of taxation. No rule that is not as wide as general social welfare will cover all cases.

PROGRESSIVE TAXATION VERSUS PROPORTIONAL TAXATION

A progressive tax is one of which the rate increases as the base upon which it is imposed increases in amount; a proportional tax is one of which the rate remains constant. In the case of the progressive income tax, the amount of the individual's net income is the base. As the amount of that income grows larger the rate of taxation is raised. Progression should not be confused with the exemption of small amounts of income or property from taxation. The income tax is not imposed on very small individual incomes, nor the inheritance tax on small estates. A small amount of personal property is usually exempted from the personal-property tax. The purpose of this exemption is twofold: in the first place, to collect small tax bills costs more than they bring into the treasury; in the second place, the exemption may be a concession to the opinion that the poor should not be compelled to contribute to direct taxes. Actually such exemption has much the same effect as progression.

Progressive taxation must not be confused with the classification of

taxable property. It is common for the states to tax some kinds of business at a higher rate than other kinds, or to levy one sort of tax on one or a few types of enterprise and a different sort on the remainder. Railroads and public utilities are often taxed on gross receipts, while manufacturers, retailers, farmers, and other businesses are subject to the taxes on the capital value of the properties they own.

Progressive income taxes are now imposed in most of the industrial countries of the world. Yet they are a comparatively recent development as compared with taxes on property and on commodities. Even at the present time there is still considerable debate concerning their justification and concerning the degree of progression.

Under the Act of 1936 the tax imposed on income by the government of the United States is as follows:

1. Net income, which may be defined roughly as gross income less necessary expenses, is the base of the tax. Income from state, county, city, etc. bonds, part of the Federal debt, gifts, and life insurance, is excluded from gross income.
2. Taxable net income for purposes of imposing the surtax or progressive rates consists of net income less the following deductions:

a. Single person	\$1000
b. Married person	2500
c. Each dependent of taxpayer	400
3. Taxable net income for the imposition of the normal or flat rate consists of net income less the deductions under (2) above, plus further deductions as follows:
 - a. Interest on bonds of the United States and of certain government corporations.
 - b. Ten per cent of the earned income. All net income up to and including \$3000 is assumed to be earned net income, but in no case will more than \$14,000 be considered earned net income.
4. The normal rate of the tax is 4 per cent.
5. The surtax rates for the lower brackets are as follows:

\$4000 taxable income or less	none
Excess over \$4000 and not over \$6000	4 per cent
Excess over \$6000 and not over \$8000	5 per cent
Excess over \$8000 and not over \$10,000	6 per cent
Excess over \$10,000 and not over \$12,000	7 per cent
Excess over \$12,000 and not over \$14,000	8 per cent
Excess over \$14,000 and not over \$16,000	9 per cent

and so on, but by irregular income intervals and by irregular increases in rates up to a tax of 75 per cent on all surtax income over \$5,000,000.

One argument for progression arises from the theory of equality of sacrifice, which we have already noticed in connection with the question of the distribution of the tax burden. That taxes may be imposed so as to cause each person to give up the same amount of satisfaction as every other, they must, it is said, be progressive. As income increases, its

marginal utility decreases in accordance with the principle of diminishing utility; conversely, when income is decreased by taxation, its marginal significance rises. But persons who have large incomes derive a relatively small satisfaction from a unit of income, whereas those of more moderate means derive a larger satisfaction from each unit of their incomes. If the tax is designed to cause all persons to contribute equally, it must deprive each of the same amount of satisfaction. But to do this it must take proportionately more from the rich than from those whose incomes are smaller.

This argument for progression is subject to several objections. In the first place, before one can accept it, he must agree that justice in taxation means equality of sacrifice. In the second place, while it may be granted that a tax of 10 per cent imposed on rich and poor alike deprives the latter of greater satisfactions than it does the former, it does not follow that the same distinction exists when the same rate of taxation is imposed on two rich men who have different incomes. Concretely, it does not follow that to take \$3000 from one who receives \$30,000 a year deprives him of any greater sum of utilities than to take \$10,000 from one who receives \$100,000. We may well doubt that the principle of diminishing utility applies with sufficient accuracy to money incomes above an amount that will purchase all reasonable luxuries to justify a system of progression. A final objection to the sacrifice theory as a justification of progression is the impossibility of ascertaining the rate at which income becomes less significant as it increases in amount.

It is also asserted that, without reference to sacrifice, ability to pay increases faster than income; that is, a very rich man can afford to pay a higher tax, just as he can afford to buy better food and clothing and live in a better house. The test is said to be purely objective. But as a matter of fact, the expenditures for consumption by recipients of large incomes do not vary as their incomes. Some are lavish and some are relatively frugal. One is led to suspect that what lies behind this assertion that ability to pay increases faster than income is either the sacrifice theory or the fact that because the rich have more income more taxes can be taken from them. The latter "justification" is not one that relies on any other principle than that of force. To show that the recipient of an income of \$250,000 *can* be compelled to pay more taxes than one who receives \$100,000 is a quite different thing from proving that he *ought* to do so according to an independently established criterion of social policy.

Still another and more practical reason for progression is the regres-

sivity of many of the taxes on commodities and the (shifted) taxes on houses in urban communities. Because the receivers of small incomes must spend a larger part of their incomes for consumers' goods, taxes on widely used commodities constitute a larger relative burden on their money incomes than on the incomes of the rich. To offset this regressivity of other taxes the taxes on money incomes should be progressive. But, obviously, this would not justify the steeply progressive rates now in effect in the United States.

It is argued also that income and inheritance taxes should be steeply progressive because large fortunes are often acquired through exploitation of natural resources which should benefit all the people, through monopoly profits, and through methods that are socially harmful. Since these fortunes are not socially justifiable, the state ought to reclaim them for the benefit of the community. Aside from the fact that the state should prevent fortunes from being acquired by antisocial means rather than seek to reclaim them after acquisition, there is a more fundamental difficulty. Taxes of necessity are imposed according to objective criteria. The law cannot discriminate clearly between fortunes that have been acquired by socially beneficial or socially harmless methods and those that have been acquired by antisocial methods.

Finally, those who believe that inequality of incomes, and especially great inequality, is socially undesirable in itself argue for steeply progressive taxation. To use taxation for the purpose of equalizing incomes is, of course, to apply to a limited extent, at least, the doctrine of "from each according to his ability, to each according to his needs." It is asserted that we cannot secure greater equality under the competitive capitalistic system through the modification of the relative shares going to wage-earners, salaried people, capitalists, and entrepreneurs without adopting measures that would decrease productivity, but we can redress the inequality that appears to be a necessary accompaniment of the capitalist system by means of taxation. Obviously the opponents of this proposal must attack the assumptions on which it rests, that there should be greater equality and that inequality cannot be adequately redressed by modification of distribution under a capitalistic system. These assumptions are considered in Chapter XXXI (Socialism).

Whatever may be the justification or lack of justification of progression, it has come to be generally applied in practically all modern countries. It appears to be a necessary accompaniment of increasing public expenditures for social services consumed chiefly by those who receive small incomes, and whether good arguments can be found to

support the use of taxation to redress inequality or not appears to be a matter of indifference. The fact remains that progressive taxation has come to be widely used for that purpose.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Taxes are compulsory payments, but the prices charged by publicly operated commercial enterprises are not compulsory. Is there any real difference in the matter of compulsion between a tax and the rate charged by a municipally owned water supply?

2. A current newspaper article states that the people of the United States spend more on automobiles each year than they do on all governmental services. It is then argued that we ought to spend more on such services as schools, parks, etc. Is this a good argument? Is it a valid reply to those who assert that the burden of taxation is too great at the present time?

3. As more and more of the private income of the people of this country is converted into socialized income by means of taxation, which class or classes do you suppose are likely to find that their real incomes are increased: the rich? the poorest class? the lower middle class?

4. To take large amounts by taxation from the rich to increase public expenditures is said to be disguised socialism. In which respects, if any, does the heavy taxation of the rich receive support from the doctrines of socialism? In which does it not receive support?

5. A great deal of revenue could be raised from a tax on a vital necessity of life, such as salt. Would a tax on salt be a desirable one to introduce into the United States? If not, which of the tests of a good tax does it fail to meet?

6. In some states the motor-vehicle tax varies roughly with the value of the car. Is this "just"? If so, can you justify the tax on the theory of benefit? on the theory of ability to pay?

7. "Taxes which are unequal are likely to be shifted. The progressive income tax is an unequal tax. Therefore it is shifted." Criticize.

8. An inflation is sometimes said to be a disguised form of taxation. Is this statement correct? How might a government benefit financially from inflation? Which economic groups would bear the burden of this "benefit"?

BOOK FIVE



INTERNATIONAL ECONOMIC
RELATIONS

CHAPTER XXXIII · Foreign Exchange

An earlier chapter explained the determination of the value of money in terms of commodities in general. There we were concerned with the problem of the general price level. In this chapter we shall consider the conditions that regulate the value of money of one country in terms of the moneys of other countries. Concretely, we wish to discover how the value of American money in terms of British or French money is determined. We wish to know this because American money is not usually acceptable in foreign countries. If an importer in New York wishes to buy goods in London or Paris, he must know how many dollars the goods will cost. But the sellers will not accept payment in dollars in New York, or at least they would not wish to do so were there no arrangements by which they could convert these dollars into pounds or francs in London or Paris.

THE MEANING OF FOREIGN EXCHANGE

The term "foreign-exchange rate" means the rate at which the money of one country can be exchanged for the money of other countries. Dealers in foreign exchange are engaged in buying and selling the circulating media of various countries. An American traveler about to embark for England wishes to exchange his dollars for pounds sterling, the British monetary unit, for, of course, American money does not circulate in England. How many dollars must he give for a pound sterling? If both countries are on a gold standard, as they were before 1914 and before England went off gold in 1931, each kind of money represents a certain quantity of gold. The dollar then contained 23.22 grains of fine gold, and the pound sterling contained 112.99 grains of fine gold. Since the gold content of the pound was 4.866 times greater than that of the dollar, the American could pay \$4.866 for a pound sterling without affecting the quantity of fine gold which he possessed. When two kinds of money exchanged at a rate corresponding to their gold content, the rate of exchange was said to be at "par" or at "specie parity."

When two countries are on a paper standard, the rates at which the

money of one country exchanges for that of the other country have no relation to gold, but are determined by the relative demand situations in each of the two countries for the currency of the other.

FORMS OF FOREIGN EXCHANGE

If the circulating media of all countries consisted exclusively of money, foreign exchange would be merely a matter of exchanging the money of one country for the money of another country. In reality the media of exchange in modern countries consist in very large part of "deposit currency." In speaking of deposit currency in previous chapters we have had in mind chiefly the demand deposits of private individuals and business firms. So far as foreign exchange is concerned, the deposits of *banks* in foreign banks is of far greater significance, since private individuals rarely have deposits in foreign banks. "Bank drafts" are in reality checks written by a bank against its deposit in another bank. If one goes to a bank to buy "foreign exchange," one very likely will get a bank draft or banker's bill, as it is often called, drawn on a foreign bank and calling for the payment of a certain amount of foreign money. Other forms of foreign exchange, such as travelers' checks, letters of credit, and cable transfers, are in reality equivalent to bank drafts, since they are all drawn against deposits maintained by the selling bank or its agent in some foreign bank.

Other forms of credit instruments that are used in foreign trade are commercial bills of exchange. These are drafts drawn by the seller against the buyer for the amount of the purchase and are in no way different from the domestic bills of exchange discussed in previous chapters, with the exception that they are drawn in terms of foreign money. They constitute an order on the part of the maker of the bill to the person or firm upon which it is drawn, to pay the sum of money designated. The buyer of the goods, however, may arrange that the bill shall be drawn not upon himself, but upon a bank with which he has made suitable arrangements. A draft or bill of exchange drawn upon a well-known bank obviously has an advantage over a bill drawn upon a relatively unknown businessman or firm.

FOREIGN-EXCHANGE PRACTICES

As a rule American exporters draw bills of exchange against the foreign buyers (or banks designated by the buyers). The American exporter takes this bill of exchange to his bank, has it discounted, and re-

ceives a credit on his deposit account. The American bank sends the bill of exchange to its foreign correspondent bank, which in turn sends it on to the bank or firm upon which it is drawn, for collection if it is a sight draft or for "acceptance" if it is a time draft. If it is a time draft and is presented for acceptance, the person upon whom it is drawn writes the word "accepted" across the face of the bill and signs his name. The bank presenting it for acceptance may then either hold it until it matures, or sell it in the open bill market. Whichever is done, the person or bank holding the bill earns, of course, the difference between the face value of the bill and the discounted value at which it was purchased.

It is particularly important to note the effect of this transaction on the deposit account which our American bank carries with its foreign correspondent. When the American bank remitted the bill of exchange to its foreign correspondent, it was credited on the books of the foreign bank (after the bill had been sold in the open market) with the discounted value of the bill. Thus when American banks discount bills of exchange drawn on foreigners and remit them to their foreign correspondents, they are building up their foreign balances. They are therefore in a favorable position to sell bank drafts drawn against those balances.

Most of the trade between Great Britain and the United States is settled in terms of British money. In the usual case, American exporters will draw their bills in pounds sterling rather than in dollars. British exporters also usually draw their bills in terms of British money. The American importer usually instructs the British exporter to draw the bill on some London bank with which an arrangement has been made. The American importer, of course, must supply the London bank with the necessary funds with which to pay the bill when it falls due. The advantage of having the bill drawn on the London bank rather than on the American importer is that a bill so drawn can be sold to better advantage in the open bill market. When the English exporter has drawn his bill on the designated London bank, he sends it to this bank to have it "accepted." He can then readily sell it in the open bill market or discount it at his bank, which in turn will either sell it in the open market or hold it until it matures. A week or so before the bill matures the American importer will go to his bank and buy a bank draft on London, and forward it to the London bank against which the British exporter was instructed to draw his bill, thus supplying the bank with the funds needed to meet payment when the bill matures. In this manner Americans who have bought goods from British merchants fur-

nish a demand for bank drafts drawn on London banks. The American banks will be glad to sell such bank drafts in order to dispose of their foreign balances.¹

Let us sum up the foregoing briefly. The American exporter draws a bill of exchange on the foreign importer or on a designated bank. The American exporter then has this bill discounted at his bank. He is credited on his deposit account with the discounted value of the bill. The American bank then remits the bill to its London (or other foreign) correspondent, and is credited with the discounted value of the bill. Thus American exports serve to increase the foreign balances of American banks. On the other hand, the American importer instructs the English exporter to draw upon some London bank which by previous arrangement has agreed to lend its name. When the bill falls due he goes to his American bank and buys a bank draft to meet payment on the bill. The sale of this bank draft lowers the American banker's foreign balance. Thus American exports increase our foreign balances; imports decrease them. To put it in another way, exports give rise to a supply of foreign exchange, and imports give rise to a demand for foreign exchange. In a similar manner, the buying and selling of foreign securities creates a supply of, and a demand for, foreign exchange.

Not all trade with England is settled in sterling exchange, as has been assumed above. Before the war approximately 90 per cent was so settled. In recent years the percentage has been smaller. There is a considerable amount of the trade between the two countries which is settled in dollar exchange. In this event the British exporter draws on a New York bank, and British exports tend to pile up British balances in New York. On the other hand, the American exporter may be instructed to draw his bill in dollars on some New York bank. When the bill matures, the British importer goes to his bank and buys a draft drawn on New York. Thus, if dollar exchange is utilized, British exports give rise to a supply of bills on New York, and British imports give rise to a demand for bank drafts drawn on New York.

¹ Banks sell foreign exchange even though they carry no balances in foreign centers. They are able to do this by means of arrangements made with the large New York banks. An illustration will indicate the manner in which this is done. The writer frequently remits drafts to English booksellers. He goes to a local bank and buys a sterling demand draft. The local bank draws a draft for the desired sum directly against (say) the Midland Bank, Ltd., London. The local bank carries no balance with the Midland Bank, but it has an arrangement with the Chase National Bank in New York by which it can draw drafts against balances carried by the latter institution in the Midland Bank. The amounts thus drawn by the local bank are deducted from its balance with the Chase National Bank. In addition to the amount of the draft the customer is charged a small sum for the service.

If the exporters of each country followed the practice of drawing their bills in terms of the money of the country in which the goods were sold, the net effect would be that foreign balances would tend to pile up in both countries. Such balances could easily be disposed of by the simple device of the two correspondent banks' drawing drafts against each other.

Before the war much of the trade of the United States with continental Europe, Asia, and South America was paid for in sterling exchange. London was, and still is, the chief center for foreign transactions.

THE RATE OF EXCHANGE

Gold Parity. Consider now the effect of the trade currents on the rate of foreign exchange. If the foreign purchases made by Americans exactly equal the sales made to foreigners, the American bankers' foreign deposits will always remain constant, and the bankers can afford to continue to buy and sell exchange approximately at par, making their profit by selling at a little higher rate than the one at which they buy. But if Americans sell more to foreigners than they buy from foreigners, the dealers in exchange will be unable to get rid of their foreign balances; that is to say, they will be unable to sell enough bank drafts to keep their foreign balances constant, since every sale by American businessmen tends to build up these balances and every purchase tends to reduce them. With huge foreign balances the American banks have these alternatives: they may invest these surplus balances by purchasing stocks, bonds, or commercial paper in the foreign country, or they may have the balance transferred to their own vaults, by ordering a shipment of gold. The latter obviously involves an expense; and rather than incur this, they would prefer to sell bank drafts below the par of exchange, though not below a point which would cover the expense of shipping the gold. With huge foreign balances, exchange rates might then tend to drop to the "gold-importing point," or (in the case of the predepression British rate) to about \$4.84. Conversely, if American businessmen are buying more from foreigners than they sell to foreigners, the demand for bank drafts will be so great that the foreign balances of our banks will prove insufficient. The American bankers therefore will be compelled to ship gold to their correspondent banks, thus building up a balance which will supply the demand. But they will be unwilling to do so unless they can sell drafts at a price that will defray the cost of shipping. Thus exchange will rise to the "gold-exporting point," or to about

\$4.89 when England was still on the gold standard. Competition among dealers will normally prevent it from going higher.

When both countries are on a gold basis, the rate of exchange will be somewhere between the "gold-shipping points." Exchange rates, of course, do not jump at once from par to one of the gold-shipping points. If the trade currents and capital movements (purchase or sale of foreign securities) indicate that the foreign balances will prove excessive, the rate of exchange will gradually drop below par; if, on the other hand, they indicate a probable deficiency in the foreign balances, the rates of exchange will gradually rise.

American purchases of merchandise or services from foreigners tend, then, to raise the rate of exchange, whereas sales to foreigners tend to lower it. It must be remembered that we buy from abroad not merely merchandise imports but also shipping services, banking services, the pleasures of travel, securities, and the services of invested capital.

If one country is on a silver basis and another is on a gold basis, the par of exchange will fluctuate according to the changes in the market rates of gold and silver. In other respects the forces determining the rates of exchange are similar to those discussed above.

Purchasing-Power Parity. Let us now consider the exchange rate between two countries one of which is on a paper basis, as was true in the United States from 1862 to 1879 and as was true of most European countries during and immediately after the World War, and again in a large part of the world following the abandonment of the gold standard by England in 1931. In such a situation the exchange rates may depart very far from the specie parity. The normal rate of exchange will then be the ratio of the purchasing power of money in the two countries. If, for example, the British paper pound exchanged for twice as many goods as the American dollar, the rate of sterling exchange would be approximately \$2. This rate would be the purchasing-power parity.

Let us suppose that the British exchange rate is temporarily far below the purchasing-power parity. In this situation Americans can exchange dollars for pounds sterling, which will buy much more than the dollars so exchanged. This will stimulate imports into the United States from England. The movements of goods from England into the United States will tend to raise prices there and lower them here, and this process will continue until the price levels of the two countries correspond to the exchange ratio. On the other hand, the increase in imports will result in an increase in the demand for bank drafts on London, with a

consequent tendency on the part of sterling exchange rates to rise. In this way, if the exchange rates are out of line with the relative purchasing power of the two kinds of money, trade currents are set in motion which tend to adjust the exchange rate to the two price levels.

This reasoning is based on the assumption that the prices of goods which enter into international trade fluctuate approximately in accordance with the general wholesale-price level. This is not true, strictly speaking, but on the whole there is a tendency for them to move together. In periods of great price upheaval, however, there are almost always certain shifts in demand and production which upset to some extent the usual price relationship.¹

It is, in part, because of the interactions just noted (which operate also when both countries are on a gold basis) that the payments of reparations or war loans must in the long run result in larger exports from the countries making these payments. Let us take the case of the postwar reparation payments. To make these payments the German government was compelled to increase its revenue, in part through taxation. With the funds thus secured the Agent-General for Reparation Payments then went into the market to buy foreign bills of exchange, which were sent on to the Allied governments. The result was that Allied exchange was bid up; for example, relatively more marks were offered for a given number of francs. This meant that the foreign-exchange value of the franc tended to rise, unless counteracted by other forces. The exchange rate was forced out of line with the purchasing-power ratio of marks and francs. Francs converted into marks would buy more goods than the francs would buy in France. Germany thus tended to become a desirable place in which to buy, and exports from Germany tended to be stimulated.

Thus the reparation payments tended to make the exchange rates favorable to the Allied countries. Moreover, these payments also tended to affect the relative price levels of Germany and the Allied countries. The Germans were taxed by their government to pay the reparations. This tended to reduce the income of the German people² and to make them unable to buy as many goods as they would if they were not taxed so heavily. As a result of this reduction in their money incomes and the

¹ Compare Gustav Cassel, *Money and Foreign Exchange after 1914*.

² The Agent-General for Reparation Payments used the credits secured from taxation to buy the foreign drafts. The government deposits therefore were reduced, the effective volume of circulating media was curtailed, and so prices tended to fall unless counteracting forces intervened. Taxation for domestic expenditures would not have this effect, because the funds thus raised would be used to buy goods within the country.

consequent decline in their purchasing power, the prices of commodities tended to fall in Germany to a level below that in other countries, exchange rates considered. Therefore, from the standpoint of other countries, Germany would be a good place in which to buy. The Allied countries, moreover, would have a larger purchasing power because of the reparation payments. Prices would therefore tend to rise in these countries. In this situation the Allied countries would constitute an excellent market for German goods. As goods flowed out of Germany the *decline* of prices there would be checked, and as goods flowed into the Allied countries the *rise* in prices there would be checked. Thus prices would tend again to be brought back into line with the exchange rates. A stream of exports sufficient to maintain this equilibrium would continue as long as reparation payments were continued, unless these payments were balanced by new debts.

In point of fact, German industries and municipalities borrowed abroad so heavily during the period 1924 to 1929 that these tendencies were largely nullified. The payment of the reparations *tended* to stimulate German exports for two reasons: first, such payments tended to make German exchange rates favorable to the Allied countries; second, they tended to lower German prices. If prices were out of line with the exchange rates, exports would tend to flow out of Germany. Equilibrium of prices and exchange rates could be re-established only by a volume of German exports (in excess of normal) equivalent to the reparation payments. It should be noted, however, that these tendencies could be overcome, as already indicated, by counteracting factors. Owing to the large sale of German securities abroad there did not develop an excess of merchandise exports, in spite of the reparation payments. This situation will be explained in some detail in the following chapter.

Speculative Rates. To return to the main argument, we have still one more possible situation to consider. Let us take the exchange relations between two countries, one of which is so situated that not only is it impossible to obtain gold and ship it out but it is also impossible to ship goods out of the country to any considerable extent. During the war and for a considerable period after the war it was practically impossible, for example, to get any goods out of Germany. Suppose an American banker had a deposit account in Germany, either carried over from a previous period when trade was open, or obtained from the sale of securities or other property owned in Germany. Under the conditions named what could he do with his deposit balance? He obviously could sell it to someone else, but what could that person, in turn, do with it?

He could not demand gold and ship it to the United States ; he could not go out and buy goods to ship to this country and thus indirectly, through the purchase, shipment, and sale of these goods in the United States, convert his marks into dollars. If neither gold nor goods could be got out of Germany and shipped to the United States, there would be no possible way of getting these marks converted into dollars. Only one possible thing could be done with the deposit balance: it could be spent or invested in Germany. But at such a time very few foreigners would be willing either to go there and spend it or to risk an investment in German property when numerous uncertainties were present, such as the danger of revolution, the destruction of property values, confiscatory taxation, and continued issues of paper money. Clearly any person who bought those German marks would be gambling with the future. Yet there are always people who are willing to take great risks in the hope of great gains. So there would probably be a sale for the marks; but the price would have nothing tangible, such as gold or goods, to tie to. The exchange rate would be a purely speculative one, based on the personal judgments of the buyers and sellers as to the future of the country in question.

We may, then, have three possible situations. First, we have the situation in which there is free movement of specie between the two countries. The normal exchange rate in this case corresponds to the ratio of the metal content of the two kinds of money. Second, we have the situation in which there is free movement of goods but not of specie. In this case the normal exchange rate corresponds to the ratio of the purchasing power of the two kinds of money. Third, we have the situation in which there is free movement of neither gold nor goods. In this case there is no norm for exchange rates ; since there are no trade currents which can be set in motion to established any normal level, the exchange rates will be on a purely speculative basis.

In the first two cases mentioned the exchange rates often depart temporarily from normal, but the movements of gold or goods which follow tend to eliminate the divergence and bring the exchange rates back to normal.

PEGGING THE RATE OF EXCHANGE

When the exchange rates are unfavorable, importation of foreign goods becomes very difficult, but national interests may require that imports be maintained. A nation may then decide to regulate artificially the rate of exchange. During the last years of the World War, Great

Britain decided to "peg" the rates of exchange in New York. This means that a point was established (\$4.70) below which the rate was not allowed to fall. This could be done only by artificially creating a sufficient demand for bills to absorb the entire supply offered at this price. Great Britain undertook to purchase at \$4.70 all the bills offered, thus making it impossible for the rate to fall below this figure. The British government obtained the funds needed to make these purchases from loans granted to it by the United States Treasury.

EXCHANGE CONTROL

During the great depression a large number of countries instituted a rigorous control of foreign exchange. This was notably true of the countries of central, eastern, and southern Europe, and of Central and South America. Owing to the severe decline, during the depression, of export prices in agricultural and raw-material-producing countries like Australia and Argentina, the difficulty of exportation from certain industrial countries like Germany, and the intensified competition arising notably from the depreciation of the British pound sterling,—for one or more of these or other reasons, many countries found themselves with inadequate supplies of foreign exchange with which to meet their foreign commercial and debt obligations. There simply was not enough foreign exchange to go around. The demand for exchange at the established rate exceeded the supply.

In these circumstances, these countries, having already for the most part been denuded of their gold resources, were confronted with two alternatives: (1) They could permit the foreign-exchange value of their money to fluctuate according to the forces of supply and demand. Since the demand exceeded the supply at the former official rate, a free market would drive the rate up. This means that it would take more of the domestic money to buy a given amount of foreign exchange; in other words, the currency of the country in question would depreciate in its foreign-exchange value. (2) They could rigorously maintain the old rate of exchange, and limit, by rationing or otherwise, the amount of foreign exchange which each importer or foreign debtor could get. This latter policy is that known as exchange control.

In the case of Germany, exchange control was instituted in 1931. Importers were granted an allotment of exchange based on a percentage of the amount used in the fiscal year 1930-1931. As exchange difficulties increased, the allotment was reduced from 50 per cent to 25, 10, and

finally in June, 1934, it was restricted to a day-to-day rationing depending on the supply available. The transfer of short-term debt obligations was stopped in 1931, and on long-term debts a partial transfer moratorium was instituted in 1933, which was made complete in 1934. The short-term debts were gradually liquidated, however, by permitting the sale of these marks to American tourists at approximately 40 per cent discount. This action stimulated tourist expenditures in Germany and at the same time enabled American owners of balances held in Germany to dispose of them (though at a loss) in exchange for dollars. In like manner, when the transfer of long-term debts was stopped, "scrip" marks were issued to the foreign owners. These "scrip" marks could be sold at a discount to foreign importers of German goods. This amounted to a subsidy on German exports. Finally, in 1934, Germany began to issue "compensation" marks which were used to pay foreigners for raw materials exported to Germany. These foreigners in turn sold such marks, at whatever discount was necessary, to their own nationals who wished to import German goods. To make up for the loss incident to the discount, Germany agreed to pay such foreign exporters of raw materials a price considerably in excess of the prevailing market.

The "tourist" and "scrip" mark systems amounted to a subsidy on German exports paid at the expense of foreign holders of German debt obligations. The system of compensation marks amounted to a tax on German imports designed to subsidize German exports.

THE INTERNATIONAL GOLD AGREEMENT OF 1936 AND EXCHANGE RATES

Under the tripartite agreement between the United States, Great Britain, and France, fluctuations in exchange rates are sheltered from the disturbing effects of trade and capital movements by the purchase and sale of gold between the Exchange Stabilization Funds of the three countries. If these movements should persist, however, the exchange ratios would eventually have to be altered. Similarly the funds may bottle up the incoming gold and keep it from getting into the banking reserves, while, on the other side, outflowing gold may be taken from the holdings of the funds without drawing upon the reserve resources of the banking system. Under this arrangement, therefore, gold movements are prevented from exerting the direct influence upon domestic prices or upon exchange rates which was formerly the case.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Is there any good reason why a formal organized market, such as the stock exchange, has not been established in which foreign exchange is bought and sold?

2. Do finance bills (that is, bills drawn by banks in excess of their foreign balances, and, therefore, constituting in effect a method of borrowing from abroad) have any influence in steadying the rate of exchange? Explain.

3. Some bills of exchange are of such unquestioned reputation that no security in the nature of accompanying documents are necessary. Other bills, however, as *documentary bills*, have certain documents attached, such as a bill of lading, an invoice, and a certificate of insurance. These documents, without which the importer cannot obtain possession of the goods, will not be released until the bill has either been paid or "accepted." In what respect, then, are documentary bills superior to "clean bills," which have no documents attached?

4. Indicate how each of the following would tend to affect the rate of exchange in New York on London: exports of merchandise; American tourists traveling abroad; sale of American securities abroad; payment by the British government of a part of its debt to the United States; payments of interest on foreign bonds owned by Americans; payment by Americans to foreign shipping companies.

CHAPTER XXXIV · The Balance of Trade



It is a common error to suppose that it is especially desirable for a country to export more than it imports. This is an ancient fallacy. The mercantilists considered an excess of exports to be a "favorable balance of trade." We still use the term "favorable balance of trade" to indicate an excess of merchandise exports and "unfavorable balance of trade" to indicate an excess of merchandise imports, but we no longer believe that the terms are well chosen. We do not admit that either one is necessarily more favorable than the other; rather, every country tends to pass through several stages in its development, and in certain periods a favorable balance of trade, so called, is the normal thing, whereas in others an unfavorable balance naturally will obtain. In order to escape this unfortunate terminology, an excess of exports is sometimes called an *active* balance and an excess of imports a *passive* balance.

OUR BALANCE OF TRADE BEFORE 1873

Up to 1873 the United States had an "unfavorable" balance of trade. This followed from the fact that up to this period we were a borrowing country. We had abundant natural resources and numerous opportunities for profit-making, but little accumulated capital. All industries were greatly in need of capital; consequently we were constantly selling industrial securities abroad. The state and national governments also borrowed heavily in this period, the latter particularly during and after the Civil War. The floating of these loans in European countries created American balances in foreign countries which could be used to buy foreign goods. Such purchases were stimulated by a favorable exchange rate. This was true because the sale of securities abroad tends to increase American bankers' balances in foreign countries and thus to lower the foreign-exchange rate. In other words, the dollar appreciated in value. It was therefore advantageous for Americans to convert their dollars into pounds sterling and to buy exportable merchandise in England and ship it into the United States. The sequence of events was as follows: the sale of securities abroad depressed the foreign exchange rate,

that is, raised the foreign-exchange value of the dollar. This stimulated imports into the United States. On the other hand, low exchange rates discouraged exports; for while the price received for American exports was established in the foreign markets in terms of foreign money, the American exporter received fewer dollars in exchange for foreign currency. The increase in imports and decrease in exports tended to rectify the exchanges, but with fresh borrowings the exchange rate again tended to be depressed. As a matter of fact, particularly in the decade prior to 1873, foreign exchange was almost constantly depressed, owing to the sale of American securities abroad.

A country which is borrowing heavily is therefore almost certain to have an excess of imports. The excess of imports follows from the fact that it is a borrowing country. If it is desirable for such a country to borrow capital, it can scarcely be said that the resulting excess of imports can properly be referred to as an "unfavorable" balance of trade. Yet we technically so refer to it because that has become for historical reasons the settled usage.

OUR BALANCE OF TRADE AS A DEBTOR COUNTRY FROM 1873 TO 1914

In 1873 the volume of loans placed by the United States in Europe reached a peak. Subsequent to that year the borrowing abroad rapidly declined, and by 1876 the repayment of these loans began. Moreover, there was the interest to be paid on the past borrowings. From 1874 to 1914 the United States could therefore properly be called a mature debtor country, a country which had already borrowed extensively abroad and was now in process of paying interest on past borrowings and repaying past loans. Of course it must not be inferred that in this period Americans placed no fresh loans in Europe, but the payments on past loans overshadowed the current borrowings.

The payment of interest on past foreign loans, as well as the repayment of loans, placed dollar credits in the hands of foreigners. This gave them enlarged purchasing power with which to buy American goods. Our exports therefore increased relative to our imports, and our trade balance reversed itself as compared to the period prior to 1874. Merchandise exports now exceeded merchandise imports. We had a so-called favorable balance of trade.

Labor and capital now tended to be diverted toward the exporting industries, since these industries were stimulated by the increased for-

eign demand. In a large country like the United States, with a relatively unimportant foreign trade, this probably did not result in any actual movement of capital and labor from the domestic industries to the exporting industries. A relatively larger proportion of new capital and of the growing population probably went into the exporting industries than otherwise would have been the case. Thus, from the standpoint of the exporting industries the transition period from an unfavorable balance of trade to a favorable balance is genuinely advantageous, though from the standpoint of labor and capital as a whole no such advantage is evident. And in the long run the temporary advantage enjoyed by the exporting industries soon disappears as more labor and capital are diverted into these industries. The relative profitableness of the domestic and exporting industries is thus rapidly equalized.

INCREASED EXPORTS DURING THE WORLD WAR

With the coming of the World War the United States passed into a third stage. Europeans found it to their advantage to purchase huge quantities of American exportable products which could not be obtained in sufficient quantity in Europe, or at any rate only at less favorable prices. In order to finance these purchases, large blocks of foreign securities were sold in this market and some foreign-owned American securities were resold here. To some extent the purchases were paid for by shipments of gold. American exports accordingly increased enormously. Export industries were enjoying an immense prosperity, and more labor and capital were diverted into the production of those goods which were demanded by the European nations.

THE UNITED STATES AS A CREDITOR NATION

Before the war the total American securities held abroad amounted to about five or six billion dollars; during the war American investors repurchased the greater part of these securities, thus practically wiping out our former indebtedness. Moreover, during the war the United States government lent some ten billion dollars to the Allied governments. In addition, private investors purchased during the war and postwar years some fifteen billion dollars of foreign securities. Thus, from 1926 to 1930, three quarters of a billion dollars of new foreign issues and direct foreign investments were placed in the United States annually. But this was offset, in large part, by the repayment by for-

eigners of past loans and the purchase of new securities. Thus our net long-term investment abroad averaged about half a billion dollars annually from 1926 to 1930.

It was largely because of these purchases of foreign securities by American investors that we continued in the decade of the nineteen-twenties to maintain an excess of merchandise exports over imports. Before the war the export excess amounted to approximately half a billion a year; during the war the trade balance rose rapidly and reached four billion dollars in 1919. It then fell to an annual average of three quarters of a billion from 1922 to 1930 inclusive; from 1931 to 1935, however, the excess declined to an average of less than one third of a billion a year.

Table 54 gives our balance of trade for the years 1935 and 1936. In 1935 the net excess of merchandise exports was \$236,000,000. This was more than balanced, however, by our net tourist expenditures abroad, which amounted to \$292,000,000. Other "invisible imports," including shipping and freight, immigrant remittances, and government transactions, were short of balancing the net amount due us on interest and dividends. Thus, all told, foreigners owed us \$208,000,000 on all commodity and service items combined. Put in another way, we sold to foreigners on current account \$208,000,000 (including service on capital, that is, interest and dividend payments) more than we purchased.

This relatively small amount could have been balanced by a corresponding gold import. In point of fact, however, the gold imports were enormously greater, amounting to \$1,739,000,000. This huge inflow of gold resulted from a net investment of \$1,537,000,000 by foreigners in the United States. These investments consisted partly of purchases of American securities, partly of the repatriation of their own securities, and partly of short-term funds.

In 1936 commodity exports exceeded imports by only \$34,000,000. Moreover, tourist expenditures abroad increased to nearly half a billion dollars, so that on current account the tables were turned and we owed foreigners the relatively small sum of \$132,000,000. The huge gold imports again were mainly the result of a large inflow of capital.

These facts are set forth in Table 54. This table presents a very different picture, in many respects, from similar tables for the decade of the nineteen-twenties. From 1920 to 1924 a major feature of the balance of payments was a large gold inflow, amounting to an annual average of about \$300,000,000. The export of gold to the United States plus the sale by foreigners of new securities in the American market (together

TABLE 54. Balance of Payments of the United States for 1935 and 1936, in Millions of Dollars

CREDIT			DEBIT		
<i>Commodity and Service Items</i>	1935	1936	<i>Commodity and Service Items</i>	1935	1936
Exports of merchandise	2283	2453	Imports of merchandise	2047	2419
Shipping and freight services (to foreigners)	63	65	Shipping and freight services (to Americans)	99	125
Tourist expenditures (in United States)	117	122	Tourist expenditures (abroad)	409	495
Immigrant remittances (to United States)	5	5	Immigrant remittances (abroad)	120	143
Government transactions (by foreign governments in United States)	28		Government transactions (by United States government abroad)	83	
Interest and dividends (paid to United States investors)	521	575	Interest and dividends (paid to foreigners)	146	200
Miscellaneous items (net)	95	30			
<i>Gold and Currency Movements</i>			<i>Gold and Currency Movements</i>		
Currency imports (net)		20	Gold imports (net)	1739	1030
			Silver imports (net)	336	180
			Currency imports (net)	1	
<i>Capital Movements</i>			<i>Capital Movements</i>		
Long-term investments (by foreigners in United States)	2009 ¹	792 ²	Long-term investments (by Americans abroad)	1547 ¹	
Short-term investments (net) (by foreigners in United States)	1075 ²	479 ²			
Residual item (represents balance of errors or un- knowns in estimates)	331 ²	201 ²			

with direct investments abroad) enabled foreigners to pay for the excess credit (about \$1,000,000,000 annual average) due us on the commodity and service account. From 1925 to 1929 there was, on the average, no net gold movement. The annual excess credit on current account of nearly \$500,000,000 due us in this period was balanced exclusively by the investments made by Americans in foreign countries. From 1930 to 1933, despite large temporary inflows and outflows of gold, there was, on balance, no net gold movement for the period as a whole. Again an average excess credit on commodity and service items of some \$300,000,000 was offset by an outflow of capital, at first both long-term and short-term, but after 1930 limited to short-term withdrawals of funds held by foreigners in the United States. The period from 1934 to 1936 was characterized chiefly by a great reversal in capital movements. Investment capital and short-term funds began to flow into the United States in large volume, especially in 1935 and 1936, the average

¹Includes refunding.

²Net.

for the period being nearly a billion dollars a year. In addition to supplying us with all this capital, foreigners owed us on balance about \$300,000,000 annually on commodity and service account. Foreigners discharged these obligations by sending us annually approximately \$1,350,000,000 of gold.

Summarizing the main developments, one may say that from 1920 to 1933 our excess "exports" of commodities and services were paid for partly by gold imports but mainly by large capital exports abroad. From 1934 to 1936, our excess of commodity and service "exports" was paid for by gold imports, but much more important was the huge inflow of gold incident to a large capital inflow from abroad.

The extraordinary reversal in the position of the United States with respect to international capital movements—having been a large exporter of capital from 1914 to 1933 and then suddenly becoming a large importer of capital—needs an explanation. During the depression, Americans liquidated some of their foreign holdings and made virtually no fresh investments abroad. On the other hand, after 1933 foreigners found it profitable to invest in American securities. This was due partly to the exceptional bargains which the severe deflation of security values afforded in America, partly to the favorable prospect for a rapid revival, particularly in the latter part of 1935 and 1936, and partly to the danger of war or of political upheaval in various European countries. Moreover, the extreme devaluation of the dollar in January, 1934, facilitated the inflow of gold which provided the ready funds for investment by foreigners in this market.

All this raises the question whether the United States has truly become a creditor nation, or whether the foreign lending of the war and postwar periods was merely a temporary episode. The present value of American foreign investments is probably about \$12,000,000,000, while the value of foreign holdings in the United States is about \$6,000,000,000. We received in 1935 on interest and dividend payments about \$520,000,000, while we paid to foreigners only \$150,000,000. We are therefore, in fact, still a net creditor nation, though if the recent capital inflows continue that position will gradually be wiped out. On the other side, with increasing world revival, some investment by Americans abroad is likely to be resumed, particularly in Canada and Latin America.

It is clear, however, that we have not yet reached the position of a mature creditor nation. This England has long been. Regularly she has

been a net investor abroad. But the interest and dividend receipts on her past investments overshadow the new investments. This is the certain earmark of a mature creditor country.

ECONOMIC EFFECTS OF AN UNFAVORABLE BALANCE, SO CALLED

Is an excess of imports really unfavorable, properly speaking? What is feared is that prices will be forced so low in this country that American entrepreneurs will be forced out of business; but even a brief examination of the facts will show that this view is entirely untenable. If our imports exceed our exports, as they must in the long run if we continue to be a creditor nation, it must necessarily follow that this country will be a favorable place for foreigners to sell in or they will certainly cease to export to this country. It must be true also that this country will continue to be a favorable place for our own producers to sell in or they will certainly divert their production to foreign markets. In short, our price level must remain relatively high compared with European price levels, foreign-exchange rates considered. In other words, the rates of exchange as compared with the price-level ratios must be such that the American market is sufficiently superior to the foreign markets to bring in the excess of imports over exports which our position as a creditor nation requires. The fear that American prices under the competition of these imports will drop in relation to European prices (exchange rates considered) is therefore entirely unfounded. In fact, the imports would not continue to come in were it not that the American market was slightly superior, prices and exchange rates considered; and this superiority is due to the increased purchasing power which a creditor nation necessarily has.

To put the matter in a somewhat different way: it is feared that the excess of imports will displace American labor and capital and prevent their productive employment. But if we consider the "imports" and "exports" of merchandise and services, omitting from the service item interest and dividend payments, it follows that in the *long run* the excess of such imports can only equal the payment made on loans either in the repayment of principal or in the payment of interest and dividends. If the imports exceeded this amount, the American demand for foreign exchange would cause exchange rates to rise above par and as a result imports would be checked and exports would be encouraged; in other words, the excess of imports over exports can never in the *long run*

exceed the additional purchasing power which comes into American hands because of the repayment of loans or the payment of interest and dividends. The remaining imports are paid for by the purchasing power derived from our exports, leaving as much purchasing power as before with which to purchase American-made goods.

There is, therefore, never any danger from a long-run point of view that foreigners will be able to destroy our home market by excessive exportation to this country. Nor is it true that we cannot keep our labor and capital employed unless we can dispose of our "surplus products" in foreign markets. "Surplus products," as popularly used in this connection, apparently means products which cannot be sold at home. Now in point of fact no nation ever produced goods that it did not itself purchase directly or indirectly. If a country exports more than it imports, it is always either in the process of increasing its investments abroad or it is a debtor country whose payments on interest exceed its new borrowings. If it is increasing its investments, the exports will exceed the imports as much as the new loans exceed the interest on the old loans. It is not difficult to see that such a country itself supplies the purchasing power which buys these so-called surplus products. The case is simply that the people of such a country prefer to invest part of their income rather than spend it, and the comparative rates of interest are such (exchange rates considered) that it is more profitable to invest in foreign countries than in their own country. Such an exporting country purchases its own "surplus." If the country is a debtor country, it cannot properly be said that it really has produced the exported "surplus." Rather the capital owned by foreigners in this country produced it; and, in effect, it is sold abroad to meet the payment of interest and dividends on this foreign capital. Conversely, if a country has an excess of imports, it is either a borrowing country or a creditor country. If it is a borrowing country, foreigners are really furnishing the purchasing power with which the excess is bought. Properly speaking, the foreigners are the real purchasers of this excess of imports. If the country is a creditor country, its total production is to be measured not merely by what is produced on its own soil but also by the product attributable to the capital invested in foreign countries. Every nation's purchasing power is equal to its total production. Every nation ultimately buys its own product. In reality we never sell any surplus to foreigners. If we consume less than we produce, we ourselves buy that surplus, which may be invested either at home or abroad; if we consume more than we produce, we are living off our capital. We may, however, consume more

than we produce on our own shores, the excess being produced by our capital invested abroad. That is the situation in which creditor nations generally find themselves. Creditor nations are rich nations; and from that point of view it may be said that creditor nations are the most favorably situated, yet they regularly have (unless they are making new loans which more than counterbalance the service on their past loans) an "unfavorable" balance of trade.

FACTORS GOVERNING THE BALANCE OF TRADE

Nor is the fear well founded that purchases of foreign goods, services, or securities "will drive money out of the country." If such purchases exceed our sales to foreigners, the demand for foreign exchange, it is true, may rise above the gold-shipping point (both countries being assumed to be on a gold basis). Under the automatic gold standard certain forces were set in motion which tended to check the flow of gold out of the country. In the first place, the high rate of exchange tended to stimulate exports and discourage imports and thus put a stop to the excess of purchases over sales which originally produced the tendency. In the second place, the outward flow of gold tended to reduce the bank reserves, raising temporarily the rate of interest, and thus making the country a favorable one in which to invest capital (that is, to buy securities); or, to put it another way, the high rate of interest tended to lower the prices of securities in the country in question, making it a favorable country in which to buy stocks and bonds. But such sales of securities to foreigners resulted in the drawing of a large volume of bills on foreigners, increased the supply of foreign exchange, and thus tended to lower the rate of exchange, and to stop the flow of gold. In the third place, the outward flow of gold tended to cause a contraction of the circulating media of the country, which in turn would result in a fall in prices, making the country in question a desirable place in which to buy. The consequent increase in exports gave rise to an increased supply of foreign bills, American bankers' balances abroad were increased, rates of exchange tended to fall, and the flow of gold to be checked.

Thus the factors that, under the gold standard, tended to regulate the currents of foreign trade were (1) the rate of exchange, (2) the rates of interest in the respective countries, (3) the price levels at home and abroad. If the price level were higher in one country than in another (by which we mean that the purchasing power of its money, in terms of

goods, was less than the purchasing power of the foreign money, exchange rates considered), goods would tend to flow into the country where the price level was high. But this movement of imports would itself tend to rectify the divergence between the rate of exchange and the price level. In like manner, if the rate of interest were relatively high in one country, investments would be made in that country and securities would be purchased. Such purchases would tend to rectify the divergence between the exchange rates and the normal relation of interest rates in the two countries.

Under the current domestic and international monetary arrangements, a large measure of governmental control is exercised over exchange rates, interest rates, and price levels. Accordingly the ultimate determinants of international trade and capital movements have been subjected to a more or less rigorous social control.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "Highly developed industrial nations cannot survive without an excess of merchandise exports over imports." Criticize, and show why the opposite has usually been the case.
2. "The United States is not yet a mature creditor nation." Do you agree?
3. "In the period 1934 to 1936 the United States was selling goods and securities and taking, in large part, gold in payment. This was a hazardous undertaking, since the value of gold is likely to decline greatly in the future." Give arguments for or against this statement.

CHAPTER XXXV · The Law of Comparative Advantage



It is the purpose of the present chapter to explain the fundamental principles that determine the currents of international trade. We have postponed this chapter until the principles underlying the distribution of a nation's income into wages, rent, interest, and profits have been dealt with, because this analysis is necessary for a thorough understanding of the law of comparative advantage and its effect on international trade.

In this chapter we shall consider (1) why nations engage in international trade, (2) the extent to which it is advantageous for a nation to engage in international trade, and (3) the grounds for the frequently expressed fear of foreign competition.

THE BASIC REASON FOR INTERNATIONAL TRADE

The basic reason for trade of any sort, whether domestic or international, is the advantages to be derived from industrial and geographical specialization. Nearly everyone in modern industrial societies produces for the market. Such specialization involves, of course, an enormous amount of trade; but this trading is worth while, for the specialization that it makes possible increases our total output enormously. Trade and transportation in themselves are so much economic waste except in so far as they serve to develop division of labor and territorial specialization and thereby increase the total volume of production. Thus it is often said that it is a great waste for farmers to ship their hogs and cattle to Chicago, only to have pork and beef shipped back to the rural towns. Or take the example of the Wisconsin cheese factories which sell their output to wholesalers and storage men in Chicago, where the product is cured, graded, and packed. These wholesalers sell to retailers in all parts of the country, including the small Wisconsin towns where the cheese was originally made. This looks like wasted effort, but it is really efficient division of labor. The market is so restricted in these small towns that the work of storage, packing, grading, and wholesaling cannot be done efficiently there either by the retailers as a side issue or by specialized dealers.

Trading not only makes human specialization, or division of labor, possible, but also territorial and geographical specialization. If the population is mobile it will tend to become distributed over the country according to the natural advantages of the sites and resources of the different regions. Each part of the country specializes more or less in a limited number of products. No doubt any one of these regions could, if necessary, produce a much greater variety of products, but it is not economical to do so, since they can be produced to greater advantage elsewhere.

COMPARATIVE COST IN REGIONAL TRADE

Let us suppose for the sake of simplicity that a nation's consumable goods can be classified under three heads, A, B, and C, and that the country can be divided, from the standpoint of the nature of its resources, into three regions, Nos. 1, 2, and 3. Let us suppose that any one of the three regions can produce all the three varieties of finished products named, but that region No. 1 has a comparative advantage in the production of A, the next best product being B; that region No. 2 has an advantage in the production of B, the next best product being C; and that region No. 3 can best produce C, the next best product being A. Now any one of three possible situations may exist: (1) we may find that region No. 1 produces A exclusively, that region No. 2 produces B exclusively, and that region No. 3 produces C exclusively; (2) we may find that region No. 1 produces commodities A and B, that region No. 2 produces commodities B and C, and that region No. 3 produces commodities C and A; or (3) we may find each of the three regions producing all three commodities.

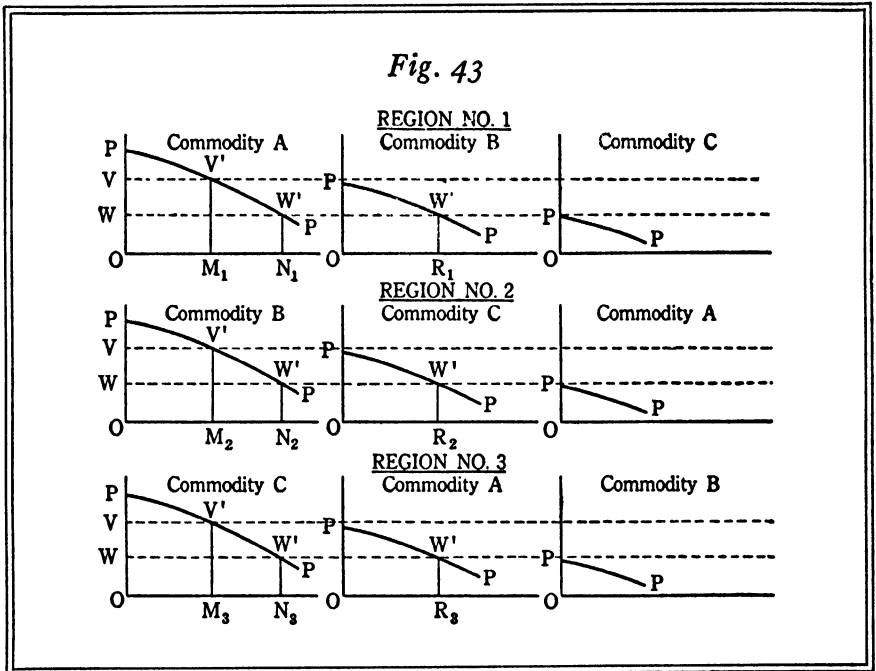
Certain very definite factors, which can best be set forth by the aid of diagrams, will determine which of the three possible contingencies will actually obtain.

Let us suppose that the line *OV* (Fig. 43) represents the marginal productivity¹ of labor² in this country. The diminishing productivity of labor in the three regions and for the three different commodities is represented by the curves *PP*. From the diagram it will be seen that *OM*₁ units of labor are employed in the production of commodity A in

¹ This term is used throughout this chapter to mean *value* productivity.

² Attention should be called to the fact that, for the sake of simplicity, we have centered our attention in this chapter exclusively on *labor*. Precisely the same analysis applies, however, to the other factors of production. We have presented the theory of international trade, not in terms of the discredited labor theory of value, but in terms that harmonize with the modern theory of value and the productivity theory of distribution.

region No. 1, that OM_2 units of labor are employed in the production of commodity B in region No. 2, and that OM_3 units of labor are employed in the production of commodity C in region No. 3. There is no more labor to be employed. The marginal product of labor is the same (OV) in all three regions. This naturally follows if we assume uniformity in skill required and perfect mobility of labor. Now, under the conditions indicated in the diagrams, it is clear that no labor could be employed in



the production of commodities B and C in region No. 1. The reason for this is that even under the most favorable combination of factors possible, a unit of labor cannot produce enough of these commodities to yield a return equal to the marginal-value product of labor employed in the production of commodity A. Likewise in region No. 2 the production of commodities A and C, even under the most favorable circumstances, is not sufficient to equal the value product of the marginal labor employed in the production of commodity B. Similarly, in region No. 3 only commodity C can economically be produced.

But now let us suppose that population increases in all three regions. The margin of production is now pushed lower down in each of the three regions. The marginal productivity of labor is now represented by the

line OW , and the number of laborers employed in region No. 1 by $ON_1 + OR_1$, the number in region No. 2 by $ON_2 + OR_2$, and the number in region No. 3 by $ON_3 + OR_3$. It will now be clear from an examination of the charts that region No. 1 will produce commodities A and B, that region No. 2 will produce commodities B and C, and that region No. 3 will produce commodities C and A. In region No. 1, for example, where in the former illustration it was uneconomical to produce commodity B, it is now possible to do so because of the fact that the margin of productivity has fallen to OW . A unit of OR_1 labor producing commodity B will now yield the same value product as a unit of ON_1 labor engaged in the production of commodity A. Similarly, in region No. 2 the marginal-value product of OR_2 labor producing commodity C is equal to the marginal-value product of ON_2 labor engaged in the production of commodity B.

While under the lower margin of productivity each region would be engaged in the production of two commodities instead of one, it should be noted, however, that each region is still specializing largely in the commodity in which it has the greatest advantage. For example, region No. 1 produces much more of commodity A than of commodity B, for it has a greater advantage in the production of A; and region No. 2 produces a larger quantity of commodity B than region No. 1, for it has an advantage over region No. 1 in the production of this commodity.

From the principles which have so far been developed it should be clear that if the population continued to increase and the marginal productivity of labor were pushed down still farther, the time would come when each region would produce all three commodities, though relatively more of the commodity in which it has an advantage. It would not have been possible, however, for each region to produce all three commodities were it not assumed that each region had a considerable degree of diversification of natural resources. The more limited a region is with respect to the versatility of its resources and its people, the more specialized it will be and the fewer products it will produce. It also follows from the principles illustrated above that the undeveloped region will (if it has trade connections with the rest of the world) be more highly specialized than it will be later in its history when it becomes more densely populated. In a new country, people naturally pursue those occupations and industries that give the greatest return. Gradually less advantageous industries are developed.

There are two factors that chiefly determine the extent to which any region will engage in diversified production: (1) the variety of its

natural resources and the versatility of its people; (2) the margin of productivity, which depends, so far as the human factors are concerned, (a) upon the size of the population relatively to the extent and richness of the natural resources and the abundance and efficiency of capital equipment, and (b) upon the training and technical efficiency of the population itself, including both manual and mental workers.

COMPARATIVE COST IN INTERNATIONAL TRADE

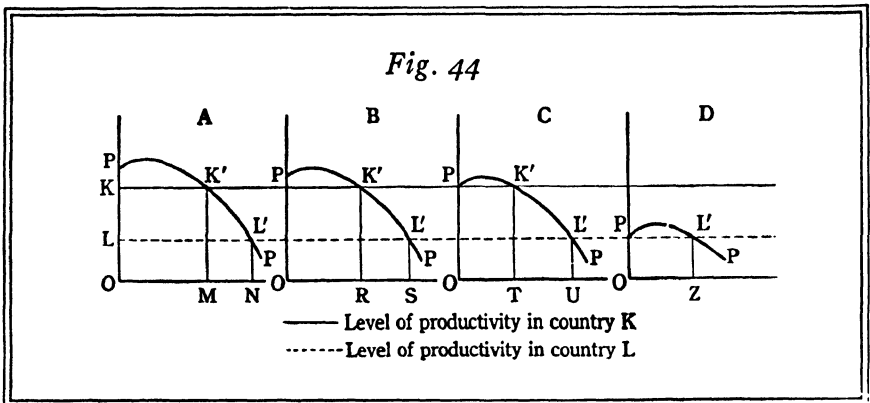
We have purposely assumed that all three regions were a part of the same country and that the population was perfectly mobile. As a result the marginal product of labor is identical in all three regions. We have done this to simplify the problem and develop certain fundamental principles which apply particularly to domestic trade and commerce but which also serve as foundation principles for the study of the law of comparative advantage in international trade. From the standpoint of the present discussion the essential difference between international trade and domestic trade is this: in international trade there are national barriers to mobility of population, some rooted deep in the traditions of race, language, and country, and others resting on legislation, such as laws restricting the free flow of immigration. Moreover, the personal efficiency of labor in one nation may be much higher than in another nation. As a result of the immobility of population and the differences of personal efficiency the marginal productivity of labor varies very greatly from one country to another.¹ Consequently many countries now produce commodities which they could not produce economically were the marginal productivity of labor uniform throughout the modern industrial world.

Let us take, by way of illustration, two countries, K and L. These two countries, let us assume, are both similarly endowed by nature with a great variety of natural resources and with equally efficient populations. But country K has a high margin of productivity and country L a low margin. The necessary principles can be illustrated more clearly by reference to Fig. 44.

In these diagrams it is assumed that both countries have exactly the same natural advantages in the production of each of the four commodities. The same set of productivity curves will therefore serve for

¹Attention should, however, be called to the fact that there is actually no such perfect mobility of labor in any country as we have assumed in this argument. In the United States, for example, there are very wide differences in wages in different sections of the country.

both countries. But the population of L is very much greater than the population of K, and as a result the natural resources of L are utilized much more intensively than those of K. Consequently the marginal productivity of labor is much lower in country L than in K. Thus, in the diagram, *OL* represents the marginal productivity of labor in country L, and *OK* the marginal productivity of labor in country K. As a result, there will be a very great difference in the relative distribution of the gainfully employed persons engaged in the production of the four commodities in these countries in spite of the fact that the natural advantages of both countries are identical. Thus *OM* persons will be employed



in the production of commodity A in country K, whereas *ON* persons will be so engaged in country L; *OR* persons will be producing commodity B in country K, whereas *OS* persons will be so engaged in country L; *OT* persons will produce commodity C in country K, whereas in country L the number so occupied will be *OU*; in country K none will be engaged in the production of commodity D, whereas *OZ* persons will be so engaged in country L. In spite of the fact that the natural advantages of both countries are identical, it is not economically possible in country K to employ labor in the production of commodity D, because the marginal productivity of labor in the case of commodities A, B, and C is too high to permit such employment; in other words, the alternative opportunities for the employment of labor in the production of A, B, and C are so good that no one can afford to employ labor in the production of commodity D. On the other hand, in country L the alternative opportunities in the employment of labor in the production of commodities A, B, and C are so poor or, in other words, the marginal productivity of labor is so low that it is economically feasible to employ

labor in the production of commodity D. Businessmen would say that it is impossible for country K to compete with country L in the production of commodity D, because of the fact that wages are higher in K than in L. This statement is perfectly true; but what is often overlooked is the fact that wages in country K are high merely because the alternative opportunities are so good that employers can pay these high wages in the production of commodities A, B, and C and still make a competitive profit.

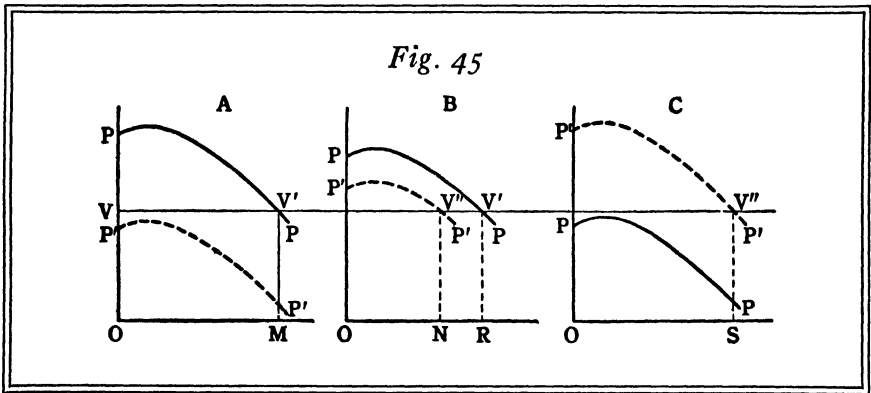
A study of the foregoing diagram should make it clear that a considerable amount of trade would be likely to take place between these two countries in spite of the fact that both were presumed to have identical natural advantages. Certainly country K will import her entire supply of commodity D. Very likely country K would import a part of her supply of commodity C, since the production of this commodity is very limited in country K and very extensive in country L. There probably would be little if any trade in commodity B, since the relative production of this commodity in the two countries corresponds roughly to their respective populations. Probably country K will export considerable quantities of commodity A to country L, since the production of this commodity in K is relatively large for her population.

From this illustration an important principle of international trade may be derived. The extent of the trade and commerce between two nations depends in part upon differences in the marginal productivity of the two countries. Since wages depend upon the marginal productivity of labor, this is merely another way of saying that the extent of trade between two countries depends in part upon the relative wage levels of the two countries. The wage level in a country may be so high that certain industries cannot be developed economically. In such a country only the industries with superior natural advantages will flourish. In these industries the productivity is so great that in spite of high wages the *labor cost* is no higher than in low-wage countries. In low-wage countries, however, even the industries with low natural advantages can be carried on profitably, for in spite of the low productivity labor costs are not excessive, because of the low wage level.

The second important reason for international trade is to be found in the advantages of one country over another in the production of certain commodities. To illustrate this principle clearly we shall assume that the wage levels of two countries are identical, but that in the production of certain commodities advantages are enjoyed by each country over the other. These advantages may be due to differences in

natural resources, nearness to markets, special training and skill of the population, or the efficiency of capital equipment. The principles involved are illustrated in Fig. 45.

The diminishing-productivity curves for country X are represented by the lines PP , and for country Y by the lines $P'P'$. The marginal productivity of labor in both countries is represented by the line OV . The number of laborers employed on the several commodities in country X is represented by OM and OR , and the number of laborers employed in country Y is represented by ON and OS . It is clear from the diagrams that country X has a great advantage over country Y in the



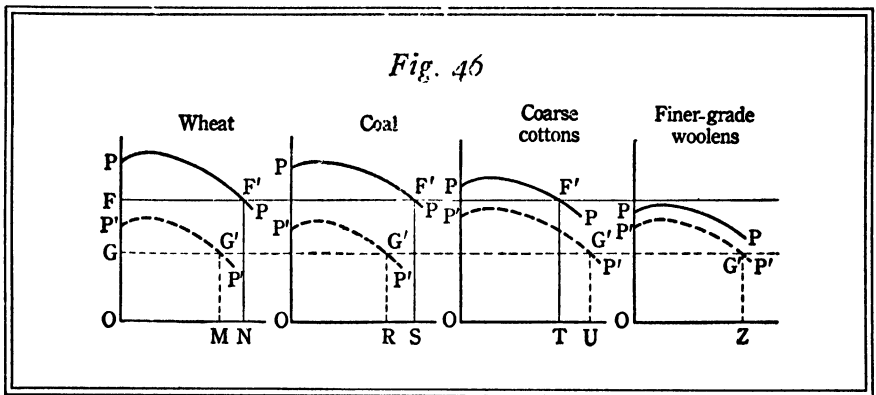
production of commodity A, that its advantage is much less in the case of commodity B, and that in the case of commodity C, country Y has a very great advantage over country X. All of X's gainfully employed population is engaged in the production of A and B, and none of its labor is employed in the production of C, since it can be more advantageously employed elsewhere. On the other hand, none of the gainfully employed population of country Y is engaged in producing A, for its labor can be more advantageously employed in producing B and C. Both countries produce B, but in each country the number of laborers so employed is less than the number engaged in the production of the commodity in which that country has the greatest advantage. It should be noted also that while both countries produce commodity B, country X has a larger number of laborers employed in the production of this commodity than country Y, because of the fact that the former country has an advantage.

In this illustration it is clear that X will export commodity A to Y and that Y will export commodity C to X. Both countries produce B, but

it is probable that Y does not produce enough to satisfy its market and therefore imports a portion of its supply from X.

These diagrams illustrate the principle that even though the productivity and wage levels are identical in two countries, trade will normally take place between the two countries because of the advantage that one country has over another in the production of certain commodities.

In the actual world no two countries have exactly the same productivity and wage levels, and nowhere do we find two countries which have the same natural advantages. Countries differ greatly, not only in the level of productivity (upon which the wage level depends) but also in



the natural advantages or disadvantages in the production of various commodities. Let us take Great Britain and the United States by way of illustration. The United States has a great natural advantage in the production of wheat and coal and, to a less extent, in the production of cloth. Let us assume that Fig. 46 illustrates the natural superiority of the United States in the production of these commodities.

The diminishing-productivity curves for the United States are represented by PP , and for Great Britain by $P'P'$. The numbers gainfully employed in the production of the various commodities in the United States are represented by ON , OS , and OT , and for Great Britain by OM , OR , OU , and OZ . The marginal productivity (and hence the wage level) for the United States is represented by OF , and for Great Britain by OG . From the productivity curves it is clear that the United States has a natural advantage over Great Britain in the production of every one of these commodities; but the advantage is greatest in the production of wheat, second in the production of coal, third in the production of the coarser cottons, and least in the case of the finer-grade woolsens.

Since the United States has a natural superiority in the production of all commodities, it may at first thought appear that she could not profitably import anything from Great Britain. It is, of course, easy to see that if Great Britain has a natural superiority to the United States in the production of certain goods it would be the normal thing for us to import those goods while we in turn would export to Great Britain the goods in which we have an advantage. But it is not so easy to see that even though we have a natural superiority over Great Britain in every line of production it will still be advantageous for us to import many things from that country. The United States cannot afford to produce everything in which she has an advantage over Great Britain.

A study of the curves will show that Great Britain will produce every one of the commodities named in spite of the fact that she is at a natural disadvantage in every one. On the other hand, the United States will produce only the three commodities in which she has the greatest advantage, in spite of the fact that she has a natural advantage in the production of all four commodities. She cannot afford to produce the finer grade of woolens at all, because the marginal productivity of labor in the production of wheat, coal, and the coarser cottons is so high that labor and capital are attracted to these fields. Labor cannot produce a value product of fine-grade woolens sufficient to equal the marginal product in the other fields. In other words, the wages that can be earned in the production of wheat, coal, and cottons are so high that no employer can afford to employ labor in the production of the finer grade of woolens.

Let us analyze this problem a little more closely. Assume that the daily wage in the United States is \$4 and that it is \$2 in Great Britain. Suppose that American workmen, under the most favorable combination of factors, can produce \$3.75 worth of woolen cloth, and that British workmen, under the most favorable combination of factors, can produce \$3.50 worth. Since American workmen can earn \$4 in other lines of production, it follows that they cannot profitably be employed in the production of woolens. On the other hand, since British workmen can earn only \$2 in other lines, it follows that they can profitably be employed in the production of woolens, and labor and capital will flow into this industry until the marginal product of labor is reduced to the general level of \$2.

It should now be clear that the United States will import her total supply of fine-grade woolens in spite of the fact that American workmen can produce more yards of woolen cloth a day than British

workmen can. On the other hand, although Great Britain will produce a part of her wheat supply, she will import most of it, since, except for the land most favorable to wheat production, her labor can be employed more profitably elsewhere. In the case of the coarser cottons the United States will import a considerable amount in spite of the fact that she has a natural advantage over Great Britain. A part of her domestic market will be supplied by her own labor—that part which can be produced under sufficiently favorable conditions to pay the wages which can be earned elsewhere.

We may now summarize the more important conclusions as follows :

1. A nation's labor and capital will naturally flow into those occupations which will yield the highest returns.

2. From the principle given above it follows as a corollary that a country may import its total supply of certain commodities even though it has a natural superiority over the competing country. Its population is fully employed in the production of goods in which it has a still greater advantage.

3. Even though a country is inferior in the production of many or even all commodities, it may still be able not only to supply its domestic market, but even to export certain commodities to the superior country. The natural inferiority in production will, indeed, mean a low wage level, but it will not prevent its capital and labor from being employed. An inferior nation need therefore have no fear that other nations with superior natural advantages will be able to undersell it in all lines and thus drive it out of business. The superiority of the foreign nations will force their employers to pay high wages, whereas the inferiority of the country in question will compel its employers to pay low wages. Thus the wage level of the superior country will be so high that the inferior country with its low wage level can compete in many lines.

4. A country with a high wage level need have no fear that low-wage countries will be able to undersell it in all lines of business. It must not be forgotten that such a country enjoys a high wage level *because* of its superiority in production. Therefore, even with the high wages its *labor costs* are no higher than those of low-wage countries. It is *labor costs*, not *wages*, that count. If it were really true that low-wage countries could undersell the high-wage countries in all lines, then, of course, it would be impossible for the high wages in the superior country to continue at their former high level, and the inevitable result would be a readjustment. But the fact that the high wage level continues is itself conclusive proof that the industries of the country can pay it.

THE MENACE OF FOREIGN COMPETITION

Why, then, do we hear so much in the business world about the fear of foreign competition if in fact these things automatically adjust themselves? The answer is that it is through the competition of entrepreneurs

that the adjustments take place. It is this competition that forces wages up or down, as the case may be, and into line with the marginal productivity of labor. The readjustments are constantly forcing certain entrepreneurs to the wall, entrepreneurs who find themselves unable to pay the going rate of wages and to sell their products at the market prices. The pressure of competition is a real one; and it is no wonder that employers are much concerned over the wages paid by their competitors, both foreign and domestic, as well as the prices at which they can sell their products.

Moreover, situations sometimes arise under which foreign competition becomes exceptionally serious. If the wage level has been raised by extraordinary circumstances in one country or lowered in another above or below the normal level established by the marginal product of labor, it is clear that the low-wage country will prove a dangerous competitor to the high-wage country. Under these circumstances the country with a wage level lower than the marginal productivity of labor would warrant low labor costs, whereas the country with an abnormally high wage level has exceptionally high labor costs. These conditions would permit the country with low labor costs to undersell the country with high labor costs. Yet this situation would set forces in motion that would tend to re-establish in each country a wage level that would correspond to the marginal product, for the country with low labor costs would capture the market and expand production, whereas the country with high labor costs, unable to get the orders, would be forced to close its shops and turn its wage-earners out on the streets. Expanding production in the former country would create a scarcity of labor, and the resulting competition for workers would gradually raise the wage level to the marginal-productivity level. On the other hand, hard times and unemployment in the high-wage country would exert a powerful pressure in the direction of a reduction of wages.

Many illustrations of such maladjustments as we have just spoken of could be cited. Often they constitute problems of major importance in modern life. Take the case of England and Germany in the postwar period. With the drop in prices in 1920 British wages failed to decline proportionately, partly because of the inertia of wages as compared to wholesale prices, partly because of social legislation, and partly because of the organized effort of the trade-unions. Consequently wages pressed severely on profits, and large numbers of workmen were unemployed. On the other hand, inflation continued in Germany until near the close of the year 1923. During this period of terrific inflation German wages

lagged far behind prices. As a result of these opposite price movements and the wage lag, British wages were abnormally high and German wages were abnormally low, abnormally, that is to say, in view of the marginal productivity of labor in each country. Labor costs were high in England and low in Germany; therefore England found it difficult to compete with Germany in the foreign markets. A more recent example is that of Japanese competition—the result of a deadly combination of modern techniques (as in the textile industry), low wages, and a greatly depreciated currency.

Similarly, if minimum-wage laws, social insurance, and collective bargains with trade-unions have the effect of raising labor costs, it is clear that the entrepreneurs in countries with lower labor standards have an advantage.¹ A larger slice of the market will consequently go to the country with lower labor costs, and the country with superior labor standards can retain them only at the expense of a certain amount of unemployment. In other words, the high labor cost will tend to force a certain number of marginal firms out of business, leaving the field to the more ably managed firms or to those having superior sites and resources. Hence countries with advanced labor standards, such as England, are necessarily interested in raising labor standards in less advanced countries in order to protect their own standards. If the standards are raised in all competing countries, a steady pressure is applied on the entrepreneurs of all countries to raise the level of productivity; but if one country alone is seeking to maintain high standards, it is likely to be worsted in the open market by the less advanced low-standard countries. Not infrequently a situation of this sort is found between different regions within the same country. Thus, in the United States the nonunion coal mines in West Virginia were able to reduce wages in periods of depression, whereas the union coal fields, working under two-year or three-year contracts, were unable to do so. In the period of depression when the market was limited the nonunion mines were able to undersell all but the better union mines. This forced the union miners to pay for their superior wages and working conditions by bearing the brunt of unemployment when the trade was depressed. This indicates the danger in raising standards higher than is justified by the industrial situation. On the other hand, if the pressure of high labor

¹ It should not be forgotten that these measures do not of necessity raise labor costs; in fact they have often lowered the labor cost through the increased productivity stimulated by establishing better conditions for the wage-earners. The increase in productivity may be due either to increased efficiency of the workmen themselves or to improved processes and methods introduced by the firm.

standards is removed, entrepreneurs are likely to seek to reduce costs by lowering wages rather than by increasing the productivity of the industry through the elimination of waste and improved management or technical processes. Competition is clearly unfair when high standards are exacted in certain regions and not in others.

Another matter that deserves brief consideration in this connection is the effect of the unequal development of economic efficiency in different countries. Before the war Germany was making exceptionally rapid strides in industrial progress. Was her growing prosperity an economic advantage or disadvantage to other countries—to England, for example? That depends. In so far as Germany's increased productivity occurred in industries in which British labor and capital were chiefly engaged, England would suffer. The enlarged output from these industries would lower the exchange value of their products, thus lowering the value product of British industries, even though the physical product was as high as before. With lower marginal productivity real wages would tend to fall until labor costs for British entrepreneurs reached the level of German labor costs. This would be the price that England would have to pay to hold her place in competition with Germany. In so far as Germany's progress developed in lines in which England was predominantly a consumer rather than a producer, England would gain, since the increased output of these goods would raise the exchange value of British industry. It is often true that a country gains from the increased productivity of its neighbors, but not always. If it is a consumer of the goods in question it gains; if it is a producer of similar goods, it may be a loser unless it can readjust its production successfully on the new basis. The doctrine of harmony of interests between nations has its limits.

It is sometimes said that a country like England will suffer in so far as Australia, India, China, Russia, and other countries develop industrially and become more self-sufficient. This is not necessarily true. It is not unlikely that as these countries expand their production their increased purchasing power will improve England's foreign market. Even though these countries do furnish in part the same types of goods and services which England offers, nevertheless they may produce so much more of other sorts of goods that, relatively, England's products will be scarcer (and therefore more valuable) than now. On the other hand, it is possible that these undeveloped countries may eventually produce so much of the goods and services which England (because of both her resources and her people) is best adapted to furnish that her

position will be weakened. The value product of a nation (and therefore a nation's income) depends in the modern order not only on its productive efficiency (and this in turn rests on (1) its natural resources, (2) its capital equipment, and (3) the efficiency and genius of its people) but also in large part on what is produced elsewhere. Any of these factors will affect England's future position in world trade.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The only reason why we gain from international trade is that the national resources and abilities of the people are different in the different countries." Do you agree? Why, or why not?

2. "The fact that the wage level is not the same in different countries is one of the leading causes of international trade." Discuss.

3. "Under free trade, goods will tend to be shipped from a low-wage country to a high-wage country until wages are equalized." Show that this is not true, and point out the reasons why.

4. "If a country has a natural or technical superiority over all other countries, it cannot afford to import that commodity." Do you agree? Why, or why not?

5. "The high wage level in Australia prevents her from developing many industries that might flourish there." Show that this is true, and point out why the wages are high. Does it follow that it is desirable for a country to have a low wage level?

6. "High wages make labor cost high in some industries but not in others." Why? Under the circumstances what will become of the high-labor-cost industries?

7. "It is an old fallacy that a nation suffers from the prosperity of its neighbors." Does a nation always profit by the prosperity of its neighbors? Does it sometimes profit? Explain.

8. "It is important that wages and working conditions be made uniform throughout the world; otherwise labor costs will be high in some countries and low in others, and this results in unequal competition." Do you agree? What fundamental point is overlooked? If wages are equal the world over, are labor costs likely to be equal? Explain.

9. "It is difficult for a nation to improve its labor standards by legislation so long as other countries do not do so." Show that this is true, and explain the reason why.

CHAPTER XXXVI · Tariff and Free-Trade Principles

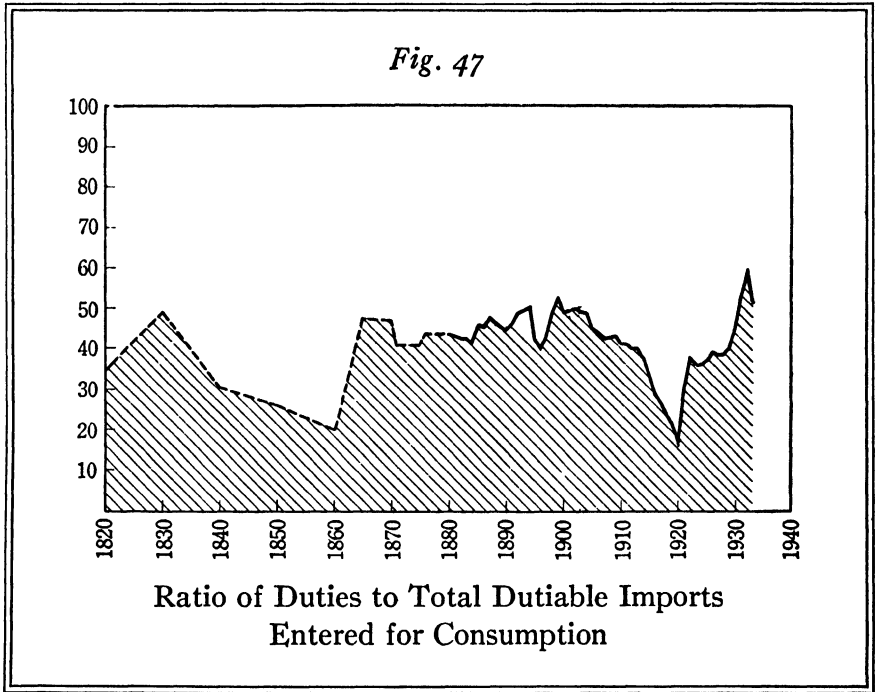


TARIFF FOR REVENUE VERSUS TARIFF FOR PROTECTION

Until recently, duties on imported articles constituted an important source of the revenue of our Federal government. However, since the import duties yielded in 1936 less than 10 per cent of the Federal revenue, it will be seen that the tariff as a producer of revenue is no longer of outstanding significance. The other function of the tariff, to protect domestic producers against foreign competition, continues to occupy a place of great importance in modern governmental policies.

In the early history of the American Union it was certainly true that the tariff was of greater significance as a producer of revenue than as a device to protect American manufacturers against foreign competition. Not until the close of the Napoleonic Wars and the War of 1812 was the tariff seriously made use of as a protective instrument. During the course of these wars commerce between the United States and England had been largely interrupted as a result of the restrictions imposed on international trade by the Berlin and Milan decrees issued by Napoleon, by the retaliatory Orders in Council issued by the British government, and finally by the embargo instituted by the United States government. This obstruction to commerce forced the Americans to rely on themselves for the supply of manufactured products to a far greater extent than before. When the war was over and the English manufacturers began to dump their surplus products on the American markets, it was inevitable that the owners of the newly formed and weak American industries should demand protection against the older and more firmly established English firms. In response to this demand the Tariff Act of 1816 was passed. In particular this act brought a considerable measure of protection to the rising textile and iron industries, the two industries that first developed the factory system in the United States. Still higher protection was granted in the Acts of 1824 and 1828. In the Act of 1828 an appeal was made to the farmers of the new West to lend their support to the protective system by including among the protected articles wool, hemp, and flax.

In the Act of 1833 the country, under the leadership of the agricultural South, which exported its surplus product of cotton and tobacco and therefore needed no protection, began to move in the direction of freer trade. This compromise Act provided that the duties should be progressively reduced until 1842. But in 1842, on the return to power of the Whigs, higher rates again were imposed. The trend toward tariff reduction returned, however, with the Act of 1846, and this was followed



in turn by a still lower tariff in the Act of 1857, when the ratio of import duties to the total value of dutiable imports fell to the level of approximately 20 per cent.

In 1861 the Republican party, which favored the policy of protection, assumed control of the government. Three tariff acts were passed during the Civil War period, each higher than the preceding one, until in the Act of 1864 the ratio of duties to taxable imports rose to nearly 50 per cent. Then began a long period of high protection during which the ratio of duties to taxable imports remained at the high level of from 40 to 50 per cent. Not until the Underwood tariff of the Wilson administration in 1913 was any material reduction achieved. Under this Act the ratio of duties to taxable imports fell to about 20 per cent in

1918-1920. With the return of the Republicans to power in 1921 a temporary tariff measure of a highly protectionist nature was passed, and this was followed by the Act of 1922, which was one of the factors that again raised the ratio of duties to taxable imports to approximately the 40 per cent level. In the Act of 1930 duties were raised still further.

The average ratio of duties to taxable imports from 1820 to 1933 is shown in Fig. 47. Except for the period of the Wilson administration, the ratio of duties to taxable imports has stood persistently at a level of from 40 to 50 per cent since the eighteen-sixties.

This chart, however, must be interpreted with caution. It is quite impossible to measure accurately the height of our tariff wall in terms of the ratio of duties to taxable imports. In the first place, it should be noted that in the case of specific duties (under which a tax consisting of a fixed absolute sum is applied per unit of imports—for example, 20 cents per square yard of cloth) the ratio of duties to taxable imports falls when the prices of imported commodities rise, and vice versa when prices decline. This factor accounts in fact for a large part of the decline in the ratio of duties to taxable imports from 1914 to 1920, and again for the great rise in this ratio from 1929 to 1932. In the second place, it should be noted that if the duties are raised so high as to become absolutely prohibitive, the ratio of duties to taxable imports, instead of rising, is likely to show a decline. This is true for the reason that commodities dutiable at a prohibitive rate will no longer be imported; therefore the ratio of duties to actual taxable imports will now be calculated on commodities taxed at a sufficiently reasonable rate to permit of importations.

In point of fact, the trend of the highly protective tariff policy of the United States, particularly as exemplified in the Acts of 1922 and 1930, was strongly in the direction of raising rates so high as to exclude all imports which could within reason, though at high cost, be produced in the United States. The result was that a larger and larger per cent of the imports actually admitted were on the free list. Indeed, before the period of high protection inaugurated during the Civil War, only 20 per cent of the imports were admitted free. Under the protective policy of virtual exclusion of the dutiable commodities, a larger and larger percentage of the total value of imports actually admitted were on the free list. From 1915 to 1933 about two thirds of the total imports admitted entered free of duty. The protective policy had become so prohibitive that for the most part only noncompetitive products, deliberately placed on the free list, could be imported.

TABLE 55. Percentage of Imports (Measured by Value) Free of Duty

	1821	1860	1900	1930
Crude materials	58.1	42.7	75.6	83.2
Crude foodstuffs and food animals		92.1	70.2	82.8
Manufactured foodstuffs		7.4	17.4	26.6
Semimanufactured goods	12.1	18.4	34.0	73.8
Finished manufactured goods	0.2	1.9	8.4	47.4

IMPORTANCE OF OUR FOREIGN TRADE¹

Before the depression about 10 per cent of all the movable goods produced on American farms and in mines and factories was shipped abroad. In exchange for these exportations there was imported a somewhat smaller value of goods from other countries. Somewhat less than 10 per cent of the goods consumed, therefore, was of foreign origin.

But the significance and importance of our foreign trade cannot be measured adequately in mere statistical terms. The utility we derive from the goods that we purchase abroad is far and away greater than the statistics on volume of imports indicate. These imports supply raw materials that are either indispensable or highly essential for our industries, and they supply wants that otherwise could not be satisfied. Were these goods rendered inaccessible, the utilities or satisfactions lost would be very great. In view of our highly protective policy, which virtually excludes all goods except those regarded as well-nigh indispensable, it is probably no exaggeration to say that no other equal volume of goods is of such great importance for our well-being as our merchandise imports.

As one surveys the range and variety of these commodities one is impressed with the vital significance of our imports for the manner of living to which we have become accustomed. We are a great agricultural country; yet we import a considerably larger volume of animal and vegetable products than we export. While we export a surplus of meat products, especially lard and pork, we bring from abroad an enormous quantity of hides and furs. We normally export grains, notably wheat, and canned, dried, and fresh fruits, but we import a much larger value of sugar, coffee, tea, cocoa, spices, and certain fruits not obtainable at home. Of inedible vegetable products we export tobacco, but we import a much larger value of rubber and vegetable oils. We export a vast quantity of raw cotton and we import an almost equal value of

¹See *International Economic Relations*, pp. 132-135. University of Minnesota Press, 1934.

silk, wool, jute, hemp, flax, and other vegetable fibers. We export a considerable supply of sawmill products, but we import a much larger value of paper and wood pulp. Of nonmetallic minerals we have a large net export of petroleum products, and we import precious stones, pottery, glass products, and asbestos. Of metals we have a net export of copper, brass, and zinc, and we are dependent upon imports for tin, nickel, ferroalloys, and a part of our lead. Even this very brief survey is sufficient to show that this vast interchange of animal, vegetable, and mineral products is vital to our economic life and the maintenance of our standard of living.

Our chemical exports and imports approximately balance, but we import very many coal-tar products, industrial chemicals, medicinal and pharmaceutical preparations, fertilizers, paints, soap, and toilet preparations, which we could not well afford to do without. Automobiles, electrical equipment, agricultural and industrial machinery we supply to all the world, and take very little of these products in return. For art works and certain types of clocks and watches, books, toys, musical instruments, and scientific apparatus we are heavily dependent upon foreign sources of supply.

In 1850 two thirds of our *exports* consisted of raw materials, and not much over one tenth were finished manufactures. In contrast, in 1930 about one half of our exports were finished manufactures, while only one quarter were raw materials. In 1850 over one half of our *imports* consisted of finished manufactures, and less than one fifth were raw materials. Conversely, in 1930 nearly one half of the imports were raw materials, whereas less than one fourth of them consisted of finished manufactures.

In 1850 our exports were mainly cotton, tobacco, and wheat. Other items were relatively insignificant. By 1930 the character of our exports had changed enormously. These three commodities were still leading exports, but with them at the top of the list now came also machinery, petroleum products, automobiles, and iron and steel products. This signifies the industrial advance of the United States and particularly the rise of the mass-production industries. In 1850 the leading imports were cotton and wool manufactures, coffee, sugar, and hides and skins. By 1930 cotton and wool manufactures, while still among the first score of imports, stood far down the list; coffee, sugar, and hides and skins, however, were still near the top. But along with these products were now added silk, paper, and rubber, each in its way an index of modern standards of living.

TABLE 56. Leading American Exports and Imports in 1935
A. EXPORTS FROM THE UNITED STATES

	VALUE IN MILLIONS OF DOLLARS	PERCENTAGE OF TOTAL EXPORTS
Cotton	390.9	17.4
Machinery	265.4	11.8
Petroleum and products	250.3	11.2
Automobiles, parts, and accessories	227.3	10.1
Tobacco (unmanufactured)	134.0	6.0
Chemicals and related products	103.1	4.6
Fruits and nuts	93.5	4.2
Iron-and-steel-mill products	88.4	3.9
Coal and coke	52.0	2.3
Copper (including ore and manufactures)	49.9	2.2
Packing-house products	43.4	1.9
Sawmill products	41.1	1.8
Cotton manufactures	38.7	1.7
Iron and steel advanced manufactures	31.2	1.4
Rubber manufactures	22.1	1.0
Paper and manufactures	20.5	.9

B. IMPORTS TO THE UNITED STATES

	VALUE IN MILLIONS OF DOLLARS	PERCENTAGE OF TOTAL IMPORTS
Coffee	136.9	6.7
Cane sugar	133.5	6.5
Rubber (crude)	119.1	5.8
Silk (raw)	95.8	4.7
Paper and manufactures	93.4	4.0
Paper base stocks and wood pulp	82.0	4.0
Vegetable oils	78.8	3.9
Tin	69.8	3.4
Chemicals and related products	68.7	3.4
Fruits and nuts	54.6	2.7
Furs and manufactures	53.2	2.6
Hides and skins	45.6	2.2
Wine and spirits	41.2	2.0
Cotton and manufactures	41.0	2.0
Petroleum and products	37.3	1.8

Increasingly our exports are taking the form of mass-production goods, such as industrial machinery, agricultural machinery, electrical machinery and apparatus, automobiles, iron and steel manufactures, refined gasoline and lubricating oils, refined copper, and automobile tires and casings. The bulk of the machinery exports go to South America, Canada, and Europe. Half of the petroleum products go to Europe, but South America, Australia, and Canada also are large buyers. Automobiles

also are sold chiefly in Europe and South America, with Canada and Australia next in rank. Iron and steel products go largely to Canada and South America. Cotton manufactures are sold chiefly in the Philippine Islands, Cuba, Canada, South America, and Central America. Rubber manufactures go largely to Canada, Europe, and South America.

The character of our imports likewise is undergoing important changes. Increasingly our imports consist of raw materials needed for modern industry and for the maintenance of our present standard of living, commodities not obtainable, or not easily obtainable, in the United States. Raw silk, coffee, crude rubber, sugar, newsprint and wood pulp, undressed furs, hides and skins, vegetable oils and fats, fruits—these are the leading imports. Raw silk comes mainly from Japan; coffee from Brazil; crude rubber from Malaya; sugar from Cuba; newsprint and wood pulp from Canada; hides and skins from Canada, Argentina, and Australia; furs from China, Canada, and Russia; vegetable oils from the Philippine Islands, Netherland India, British Malaya; tin from Malaya; and fruits from Central America.

In a significant study of imports into the United States in 1927 (see *Vierteljahrshesfte zur Konjunkturforschung*, Sonderheft 25, 1932), Professor Alfred Rühl of the University of Berlin found that some 35 per cent of our imports consist of products which either cannot or can only with great difficulty be produced at home; they include such commodities as silk, rubber, coffee, cocoa, bananas, tea, jute, copra, hemp, vegetable fibers, and diamonds. About 15 per cent consist of special quality products, such as works of art, Oriental rugs, Turkish tobacco, Egyptian cotton, fine woolens, fine cheeses, watches, and foreign books. About 35 per cent consist of products for which there is insufficient domestic production. Only 10 per cent of the imported goods, he finds, enter into direct competition with commodities for which there are adequate facilities for domestic production.

ARGUMENTS IN SUPPORT OF THE PROTECTIVE TARIFF

The Infant-Industry Argument. The arguments that have been advanced in support of the tariff have shifted from time to time in our history. Changing economic conditions have altered the arguments brought forward, and each argument advanced has been suited to the peculiar conditions of the period in question. The earliest theory advanced in support of the protective tariff was the infant-industry argument. According to this theory a protective-tariff policy is especially

valuable in a new country where the resources and people are well adapted for the industries which it is proposed to protect, but where these industries find it difficult at first to compete with the industries which are already firmly established in the older countries. The protective tariff helps to get these industries on their feet; and it was claimed that once they had become going concerns they would be able to face world competition without any sheltering governmental protection. The argument was effectively made use of after the Napoleonic Wars, when our infant manufacturing industries found it difficult to face the competition of the better-established English manufacturers.

If a protective tariff is to be successfully maintained politically, the system must be made to appeal to the general public and not merely to a specially favored class of producers; or at least it must be shown that the benefits accruing to a special class will eventually accrue directly or indirectly to other classes as well. The infant-industry argument sought to show that the building up of these industries would be beneficial to the entire nation, since eventually it would enable the domestic consumers to secure the products of these industries on more favorable terms than from the foreigners. Once the industries in question were firmly established, the whole nation, it was argued, would benefit.

The Home-Market Argument. Historically the home-market argument came next. The infant-industry argument was not sufficient to satisfy the demands of the protectionists, since it could be made to apply only to those industries which the resources of the country and the abilities of its population specially favored. It was next the desire of the protectionists to show that it was economical to develop even those industries which could never hope to face foreign competition on a free-trade basis. The home-market argument was admirably suited to this purpose.

This argument was directed in particular to the farmers, who, after the Napoleonic Wars, were suffering from inadequate foreign markets for their products. Farmers have always been hard hit by postwar deflation of prices, and the post-Napoleonic agricultural depression had many points of similarity to the agricultural depression which followed the World War. Under the leadership of Henry Clay the protectionists came forward with what appeared (superficially at least) to be a solution of the problem. And this was the proposal: American manufactures were to be built up by means of a protective tariff. The development of these industries would create a large urban population which would furnish a good domestic market for farm products. With a larger

proportion of the population engaged in manufacturing, a more favorable balance would exist between agriculture and industry. Thus it was argued that a tariff which directly protected manufacturing was indirectly beneficial to agriculture as well.

The High-Wages Argument. Until the eighteen-twenties, during which decade the wage-earners in many states gained the suffrage, it had been urged that a protective tariff was necessary because of the high level of wages in America compared with the wages paid in Europe. Manufacturing could not hope to develop, it was argued, in a country that was faced in its foreign competition with the handicap of extraordinarily high wages. But now that the wage-earners had become voters it became necessary to gain their support for the cause of protection. So the argument was shifted, and it was now made to appear that the protective tariff was the *cause* of high wages, and that a protective system was necessary to *maintain* high wages. Without this protection foreigners would be able to undersell American producers, and this in turn would either drive the latter out of business or compel them to lower wages in order to compete with the foreign-made goods. Protection would bring increased employment at high wages. Such was the argument advanced by the politicians in the campaigns of 1828, 1832, 1836, and 1840. The tariff was necessary to protect American labor against the pauper labor of Europe. And this argument has been repeated down to the present day.

The National-Defense Argument. Finally the protective tariff was urged as a means of safeguarding the interests of a nation in time of war. Wars disrupt commerce between nations and force each country to rely more or less exclusively upon its own production. It is argued that in a world which is still subject to important international conflicts every generation or so, it is not wise to depend too much upon foreigners for essential products. A certain measure of national self-sufficiency is desirable and even necessary, it is said, in a world which has not yet been able to achieve lasting international peace. This argument does not rest upon economic grounds at all, but upon purely political and nationalistic grounds.

THE VALIDITY OF THE ARGUMENTS FOR PROTECTION EXAMINED

Let us now consider briefly the validity of these arguments. The infant-industry argument theoretically is entirely sound, but practically there is great danger in it. Industries of this character are never found

willing to admit that they have arrived at manhood; they are never willing to test their strength by facing foreign competition in a free market. This has been our experience with the cotton, woolen, silk, and iron and steel industries. Having grown up under the sheltering support of the tariff, they persist in clamoring for a continuance of its protection. There are, indeed, sound theoretical arguments for protection, as the British economist Sidgwick once said, but in practice free trade is safer.

There is also much sound sense in the national-defense argument, but, as already indicated, the argument does not rest on economic grounds. Even though it were detrimental to a nation economically to build up a protective system, it might still be worth while, as Adam Smith pointed out a hundred and sixty years ago, to maintain such a policy from the standpoint of national safety. It is not economical to maintain a standing army, but it may still be necessary to do so, and the same argument may apply to the tariff.

But even the nationalistic argument is not wholly faultless. Tariffs and wars are interrelated in a vicious circle. Tariffs may be necessary as a protective device in case of wars, but the spirit of hostility toward foreign countries is fostered and inflamed by the erection of tariff barriers. A free-trade policy the world over would go far toward establishing international good will.

The home-market argument is essentially unsound except in so far as it is linked with the nationalistic argument. The home-market argument overlooks several essential facts. If indeed there is a maladjustment, a lack of balance between agriculture and manufacturing to the disadvantage of agriculture, it follows that labor and capital will gradually be diverted from agriculture into manufacturing, even though no protection be given to manufacturing. Under a protective system this movement would doubtless be more rapid; but it might easily go too far, diverting labor and capital away from farms which from the standpoint of the best economic utilization of the nation's resources ought not to be abandoned.

If, on the other hand, the protection to manufacturing results in a diversion of labor or capital or both from foreign countries to the protected country,¹ the result varies with the particular circumstances. If the immigration of foreign labor is stimulated by the tariff policy (which

¹ It is well known that foreign manufacturers frequently "jump" the tariff wall by building factories within the borders of the protected country. Thus the Ford Motor Company has built several plants in various foreign countries. The International Harvester Company and many others have pursued a similar policy.

is doubtful), the foreign market for farm products is likely to be reduced at the same time that the home market is increased, though not necessarily to the same extent. It is not impossible, however, that the foreign population may remain substantially as large as it would have been had there been no emigration. Under these circumstances the population of the world would be greater to the extent that emigration was encouraged by the protective tariff, and world-wide increase in population is, of course, favorable to the farming class or, at any rate, to the farm-owning class. Thus a protective policy might actually benefit farmers, if indeed such a policy really had the effect of stimulating the growth of population in the undeveloped country. And probably a rapid expansion of population in a new country with undeveloped natural resources is, all things considered, beneficial to the nation as a whole. It is true that the wage-earning class is not benefited by such a policy; but since in an undeveloped country the independent proprietors, whether farmers, merchants, or manufacturers, are most numerous, the general interest is in favor of an increase in population.

It is seriously to be doubted, however, that a tariff policy will stimulate the immigration of foreign labor except in the case of skilled labor, which, without the tariff policy, might not find suitable employment. If the population is not materially increased as a result of the protective tariff, the farmers will lose and not gain from this policy. Their home market will then in point of fact not be increased. Instead, labor will be drawn from agriculture into industry, and this will raise the money wages paid for farm labor. Moreover, farmers will be compelled to pay higher prices for the manufactured products, since the domestic manufacturers cannot produce as cheaply as the foreign manufacturers. If the domestic manufacturing costs were not in fact higher, no protective tariff would be needed. Thus the farmers' costs are raised all round as a result of the policy of protection. Moreover, if the tariff policy resulted in no material increase in immigration and population, the farmers would not only lose because of higher costs but also because of a restricted market. If we make our own nation more self-sufficient we thereby force foreign nations to become more self-sufficient also. If they cannot sell their products abroad they are compelled to turn their attention to supplying the home market. As a result, agricultural production is stimulated in foreign countries, and so the foreign markets for our farmers are injured. Exporting industries regularly find that their foreign markets are curtailed as a result of tariff barriers and the consequent tendency toward national self-sufficiency. The markets for

export industries are better under conditions of free trade. Hence the home-market policy fails to benefit the farmer, first because it is not in point of fact likely to develop the home population (and therefore the home market), second, because it reduces the foreign market, and third, because it raises the money costs of farm labor and manufactured goods purchased by the farmer. Thus the industrial tariff policy tends to lower the prices of farm products and on the other hand to raise the farmer's cost of living and expenses of production. And so far as an agricultural tariff policy is concerned, until the United States becomes a net importer of domestically produced farm products, a tariff on *agricultural* commodities is of no aid to the farming population.

Consider now the high-wages argument. If the industrial tariff does not aid the farmer in a country like the United States, does it not at any rate benefit the wage-earners? In answering this question there are two possible assumptions. First let us assume that the country is largely undeveloped and that the extractive industries—agriculture, mining, and lumbering—are sufficiently superior to manufacturing (considering the resources, population, and development of the country) so that without the artificial support of the tariff the entire labor force would be employed in these three industries. It is assumed, then, that agriculture, mining, and lumbering, being highly productive, are capable of paying such high wages that manufacturing cannot flourish, for it cannot pay these wages and compete in the markets of the world. Under these circumstances manufacturing cannot develop at all without a protective tariff. Assume, now, that a protective system is introduced. The prices of imported manufactured goods are now increased as a result of the tariff. This is obviously an injury to the producers in agriculture, mining, and lumbering who buy these commodities. But under the stimulus of these high prices it is now possible for manufacturing to pay as high money wages as can be paid in agriculture, mining, and lumbering, and even to pay wages enough higher to draw labor away from these industries. Thus manufacturing develops at the expense of agriculture, lumbering, and mining. Money wages are raised, but what about real wages? The tariff policy raises the cost of manufactured goods,¹ and in so far as these come within the scope of the wage-earner's

¹ The prices of the protected goods, of course, tend to rise in relation to other commodities. Moreover, the general price level tends to be high in a high-tariff country. When the tariff is raised, imports are curtailed, but exports continue at the old level for a time. The excess of exports lowers the rate of exchange and causes an importation of gold. This tends to increase the volume of credit in the protected country and therefore to raise prices. Higher prices check a further excess of exports and so operate to check any further inflow of gold.

budget his cost of living is increased. Will the cost of living rise more than money wages? The answer is Yes, because at bottom real wages depend upon the value of the productivity of the marginal laborer, and this is lowered by the tariff policy. The proof of this statement is not difficult. The real wages of labor in any nation depend finally upon the *quantity* of goods and services which can be purchased with the marginal product of labor in any given industry. The *quantity* of goods and services depends upon the physical productivity in all the various industries taken as a whole. Now will the physical productivity in industry as a whole be greater under conditions of free trade or under conditions of a protective tariff? Surely, under free trade. The physical quantity of goods and services which can be bought with the physical quantity produced by the marginal laborer in manufacturing is not equal, under free-trade conditions, to the physical quantities of these goods and services which may be purchased with the physical product of the marginal laborer in agriculture, mining, or lumbering. When labor engaged in these industries is artificially diverted to manufacturing, a loss in the purchasing power of a day's labor is sustained. By producing exclusively mineral, lumber, and farm products and buying manufactured products abroad, the nation will have a greater total available quantity of all sorts of goods than when the imports of manufactured goods are prohibited and labor which is effectively employed in mining, lumbering, and agriculture has to be transferred to the less effective industry, namely, manufacturing. The physical productivity of labor would be lower under a protective-tariff policy, and so the physical quantity of goods of all sorts that can be purchased with the products of the marginal labor in any industry will be reduced. Thus the real wages of labor are lowered as a result of the tariff policy.

But while this argument applies to labor as a class, it does not necessarily apply to all grades and kinds of labor. Skilled labor is likely to benefit from a protective-tariff policy in a situation such as the one just described. We have assumed that without the tariff policy only agriculture, mining, and lumbering would flourish. None of these industries require a large amount of skilled labor. In consequence of the small number of skilled jobs to be filled, wages of high-grade labor would very likely be low relative to the wages of unskilled and semiskilled labor. Under a protective-tariff system manufacturing would develop, and the demand for skilled labor would increase relative to the demand for common labor. Under these circumstances it is not unlikely that a protective tariff would raise the real wages of the skilled workman. Thus,

while the total national income is reduced by a tariff policy, it may still be true that under certain circumstances some of the agents of production actually benefit by a tariff policy.

Now let us consider the second possible condition. Assume that the resources and population of the country are such that not only agriculture, mining, and lumbering will flourish but also a great many forms of manufacturing, even though there be no protective tariff. Such in point of fact is the case in the United States at the present time. Among the numerous products that we can produce more cheaply than other countries, and which therefore need no protection, are automobiles, factory machinery, farm machinery, electrical appliances, sewing machines, tools and hardware, office equipment, typewriters, and certain packing-house products. These are amply sufficient to furnish a large field for skilled labor under free-trade conditions. Under these circumstances even skilled labor will very likely suffer a decline in real wages under a protective-tariff policy. The argument in support of this conclusion is precisely the same as the one developed above with respect to common and semiskilled labor.

We therefore conclude that in countries whose resources and people are adapted to develop many types of manufacturing on a free-trade basis, even skilled labor and technical experts may lose from a tariff policy, because, although the protective system will tend to develop certain kinds of manufactures requiring a high degree of skilled and technical knowledge, it is likely on the other hand to hamper other types of manufacturing (requiring an equal degree of skill) which flourish best on the free-trade basis.

INDUSTRIES INJURED BY A TARIFF POLICY

It is admitted, of course, that the protected industries benefit from the tariff. These benefits are easily seen and are likely to impress the public mind. What is likely to be overlooked is the fact that while certain industries benefit from the tariff, others are positively injured. This aspect of the matter is not so easily seen, and this explains the public's ready acceptance of the tariff fallacy. In the United States, for example, it is safe to say that many more industries are injured by the tariff than are benefited by it. All exporting industries suffer from the tariff. On the one side, it reduces their foreign markets, since the effect of a high-tariff policy is necessarily to restrict international trade; on the other side, the tariff raises their costs of production. Moreover,

all the domestic industries which have no foreign competition, and the products of which therefore are not protected, are injured by the tariff policy because their costs are raised by the tariff. Money wages (though not real wages) and the prices of the protected goods are higher as a result of the tariff, and so the costs to the nonprotected industries are raised.

Let us enumerate some of the industries that are injured by the tariff. In Table 57 they are classified under the two main heads of exporting industries and domestic industries. The list is by no means complete, but it illustrates the more important industries that are adversely affected.

TABLE 57. Leading Industries Injured by Protective Tariff in the United States¹

A. Exporting industries

1. Agriculture
2. Meat-packing
3. Agricultural machinery
4. Industrial machinery
5. Electrical apparatus
6. Office equipment
7. Automobiles
8. Petroleum
9. Coal-mining

B. Domestic industries

1. Railroad transportation
2. Municipal utilities
3. Lake and river transportation
4. Building industry
5. Trade (wholesale and retail)
6. Hotels and restaurants
7. Bakery and confectionery
8. Publishing and printing
9. Educational institutions
10. Government service

It is frequently said that the United States is a splendid example of the beneficial results of a protective tariff, that American prosperity has been built up on the basis of a high-tariff policy. In point of fact it is much nearer the truth to say that the United States is an outstanding example of the benefits of free trade. Our domestic market is the greatest free-trade market in the world. There are no tariff barriers between the states constituting the American Union. Free trade between the states has opened up a vast national market, and this has made possible mass production, intensive division of labor, and the use of elaborate machinery. It is the large-scale machine system of production which is to a large extent responsible for our high standard of living and our high wage level. Other factors are our rich and abundant natural resources, the high efficiency of our people, the efficiency of management, and the abundance of capital. Over 90 per cent of our total national

¹ Compare B. M. Anderson, Jr., "The Tariff in an Unbalanced World," *Chase Economic Bulletin*, November 19, 1923.

production goes into our domestic markets. Free trade between the states is an important and fundamental basis of American prosperity.

BRITISH AND AMERICAN TARIFF POLICIES

Why have British manufacturers until recently favored free trade, whereas American manufacturers have consistently favored a protective tariff? If we examine the problem historically the answer is not difficult to find. When the British manufacturing industries first developed, they had no competitors of significance in any part of the world. British manufacturers therefore had nothing to fear so far as the home market was concerned. Moreover, the superiority of England's manufacturing system soon developed a considerable demand for her manufactured products in various parts of the world. Her growing manufactures required the importation of raw materials for her factories and foodstuffs for her growing factory population. Britain's extractive industries, including her agriculture, faced foreign competition, however. British agriculture already had a tariff system which protected farm products. Under these circumstances the agricultural interests very naturally demanded a continuance of protection, but the manufacturers demanded free trade. In this struggle the manufacturing interests won. The victory injured the landlords, but the net effect was favorable to the general welfare of the people as a whole. As is almost always true, a program of economic reform is likely to injure certain classes, even though the general good is furthered thereby.

The American situation was quite the opposite. Here the agricultural and extractive industries exported a large part of their products and had nothing to fear from foreign competition. The manufacturers, however, faced foreign competition in the domestic market. Thus the manufacturers demanded protection, but the interests of agriculture and the extractive industries were linked with free trade. So long as the agricultural South dominated the government, the tariff accordingly was used chiefly for revenue purposes and not primarily for protection. But during the Civil War the manufacturing North acquired control of the government and has retained this control, for the most part, down to the present time. The farmers of the North—the new West—were linked with the manufacturing East in their opposition to the slave aristocracy of the South. Moreover, the farmers and the wage-earners of the North were adroitly won to the protective-tariff policy by the home-market and high-wages arguments.

It was probably inevitable, in a century in which inventions, machine processes, and the factory system were rapidly developing, that the rising manufacturing class should achieve a dominating position in industry and politics. In point of fact they became the dominating influence in England during the eighteen-thirties and in the United States during the sixties. In England the interests of the manufacturing class were favorable to free trade; in the United States they were favorable to protection. Hence the divergent tariff policies of the two countries.

But the conditions are changing rapidly in the United States. Many of our manufacturing industries now not only supply the domestic market but export large quantities besides. They are able to compete in the markets of the world, and gain nothing from the protective policy. In consequence a growing body of manufacturing interests favors the fostering of foreign trade and the reduction of trade barriers.

Yet tradition is a powerful force. It will take a long time to change ingrained habits of thought. It is possible that the bulk of our manufacturing industries even now are injured by the tariff. Yet the majority of American manufacturers still stand solidly behind the protective principle. The American public has long been schooled in the doctrines of protection, and it will be difficult to change the established tradition. There are, however, already signs of a changing attitude, notably in the public support accorded the trade agreements program.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Referring to the future revival of European industry and the program of tariff revision, Dr. B. M. Anderson, Jr., of the Chase National Bank, New York, writes as follows: "This would involve a shifting of our industries: Those lines of manufacturing production, in particular, which use a great deal of labor in comparison with the capital which they employ, will probably be obliged to give up part of their labor and to contract their operations. . . . We shall succeed best in those lines where labor is economized, and land and capital used more lavishly. Labor will be released in the United States from some lines of manufacturing, will shift back to agriculture, to the building trades, to copper mining, and to other lines which will be stimulated by the restoration of Europe or by a lower level of costs. This process should, however, involve no general slump, and should rather be accompanied by a general stimulus to business, based on the revival of certain lines which have been unduly depressed. Among these will be copper, zinc, the packing industry, the hide and leather industry, the fertilizer industry, and the farm implement industry, as well as the growing of grain and livestock." Why would free trade stimulate industries that employ relatively more capital and land and less labor?

2. A tariff must necessarily lower the total national real income. Certain factors of production will gain from a tariff policy; others may lose. Thus farmers would gain from free trade, but it is not impossible that certain types of skilled labor may lose. Defend each of these statements.

3. "Since free trade would be particularly injurious to those industries which employ much labor and relatively little capital and land, it follows that labor as a factor in production would tend to be squeezed out if the tariff were removed. It would therefore be less significant in the productive process compared with the other factors of production. Landowners and investors would gain from free trade while labor would lose." Evaluate this argument carefully. If it is admitted that labor would be worse off relatively, might it still be better off absolutely? Why?

4. Is there any ground for the conclusion that free trade might have the effect of raising temporarily the rate of interest? If the rate of interest were raised, is it likely that more capital would be accumulated? Why? What would be the effect of this on the total national income? on real wages?

5. "Protective tariffs on manufactured goods do affect the farmers adversely in two ways. On the one hand, they raise the prices of commodities which the farmers must buy. On the other hand, by checking the imports of goods from Europe, they lessen the supply of dollars available for European purchasers of American farm products and so cut under the market for our farm products and lower their prices. The farmer is on strong ground when he demands the reduction of the tariff on manufactured commodities."—B. M. ANDERSON, JR. Show that this argument is sound. How do you account for the fact that the American farmers have been willing to support the tariff policy?

6. "The tariff builds up a home market for farm products and hence tends to raise the prices of these commodities." Point out the fallacy in this argument.

7. It is true that if Great Britain adopted a protective tariff certain industries would be stimulated, but what is likely to be overlooked is that her export industries would be injured by such a policy. Show why this would be the case.

8. It is clear that if Great Britain adopted a tariff policy her labor, capital, and land would be employed in a different manner from that prevailing under free trade. The question is, Would her people and resources be employed more favorably or less favorably? What is the answer?

9. Some industries which are protected by low tariffs are really injured more than they are helped, by reason of the additions to their costs made by tariffs benefiting other people. The steel industry is probably in this class.

10. It is argued that the table on page 647 does not show a progress toward free trade but a progress toward higher protection, because it shows that comparatively little of dutiable goods was imported, since the duties on dutiable goods are so high as almost wholly to shut off imports. Evaluate this argument.

CHAPTER XXXVII · The Trade Agreements Program

POSTWAR AMERICAN COMMERCIAL POLICY¹

After a period of relatively low tariffs, the United States introduced during the Civil War a highly protective system. Except for the low tariff of the Wilson administration, which, however, was ineffective by reason of the war, the policy of high tariffs has been persistently adhered to. The period of relatively free trade prior to the eighteenth-sixties was particularly suitable to the essentially agricultural character of our economy and to our position as a borrowing country. The high-tariff policy, beginning with 1861, was designed to foster the development of manufacturing. A decade or so after the inauguration of the system of high protection, the growing interest payments due on past borrowings from abroad began to overshadow our new borrowings. In her international financial relations the United States had entered the stage of a mature debtor country. Large annual interest charges were due on the foreign capital invested here. These interest payments were paralleled more or less closely by the excess of merchandise exports which became the normal characteristic of our foreign trade during the forty years preceding the World War.

The tradition of protectionism and of an excess of merchandise exports had become deeply ingrained in the minds of the American public. Had the World War not come, the inevitable transition from a debtor country to a creditor country (one having a net inflow of interest and dividend payments on foreign investments), and the day-to-day adjustment toward an excess of imports of goods and services,² would have come piecemeal and by slow gradations. In consequence, no violent public reaction would have been likely to be registered against the inevitable emphasis on imports.

The war, however, thrust us into a creditor position almost overnight.

¹ Compare *International Economic Relations*, pp. 118-125. University of Minnesota Press, 1934.

² Under the term "services," as here used, interest and dividend payments on foreign investments are excluded.

It could not reasonably be expected that an import psychology could be developed at once. Social institutions change slowly and, even under normal processes of evolution, tend to lag behind fundamental changes in the basic technological and economic data. Taking a realistic view of social forces, one can scarcely escape the conclusion that the abrupt shift from a debtor to a creditor position could have had no other outcome than an irreconcilable conflict between hard economic realities and our prevailing political psychology toward international economic relations. In consequence we reached, perhaps inevitably, the impasse to which the policies of the decade of the nineteen-twenties brought us. The tradition of prohibition or restriction of imports, and of emphasis on exports, continued.

Moreover, the postwar business psychology and sales methods favored the sale of goods on credit. Internally, installment selling achieved a quite new position in sales technique. Real wages did not rise in due proportion to increases in productivity. The total wages and salaries paid out in manufacturing, public utilities, and mining were \$21,310,000,000 in 1923 and increased to only \$22,837,000,000 in 1929, and total wages and salaries in all economic activities increased from \$42,893,000,000 in 1923 to only \$52,793,000,000 in 1929.¹ On the other hand, the figures of the Federal Reserve Bank of New York for about five hundred corporations show an increase of net profits from \$2,480,000,000 in 1926 to \$3,347,000,000 in 1929, while the net profits of all corporations increased from \$6,640,000,000 in 1923 to \$9,130,000,000 in 1929.² Higher wages and lower profits (which a more flexible cost-price structure would have provided), instead of the extravagant program of installment selling, would have contributed to the maintenance of business stability in this decade.³ And just as installment selling in the domestic field was substituted for an appropriate increase in wages, so also foreign lending in the international sphere was substituted for

¹See *National Income, 1929-1932*, Senate Document 124, 73 Congress, 2 session; and W. I. King, *The National Income and its Purchasing Power*.

²See Statistical Abstract, 1933, p. 272; and Bulletin 50 of the National Bureau of Economic Research, p. 2. The figure for all corporations excludes tax-exempt corporations and life-insurance companies. There is some evidence which indicates that the increase in profits was, in part, a rebound from the relatively low (compared with prewar) level of profits in 1919-1923.

³This statement should not be interpreted to mean that we regard the so-called "purchasing-power theory" of the business cycle as sound. For a discussion of this matter, see Chapter XXI ("Business Cycles"). It is not the distribution of income as such, but the violent *change* in the distribution of income, particularly the huge growth of profits, which causes overinvestment and maldistribution in the capital market. From the standpoint of business stability, what is wanted is a condition of equilibrium in the distribution of income.

the enlargement of imports implicit in our new creditor position. The lack of balance, both internally and externally, was covered up temporarily by the creation of debt both at home and abroad.

Obviously, there was no conscious planning leading to this dangerous solution. International trade and international financial transactions seldom flowed through the same channels. But vast capital exports nevertheless did sustain an expanded foreign market for American goods. It was an artificial market, and our production plans consequently were based on illusory data. They served to distort the judgment of individual producers of different kinds of goods as to the true eventual market. A combination of circumstances, rather than conscious planning, accounts for the fact that capital exports bridged the gap created by the coexistence of a creditor status and an export surplus. The vast gold stocks and the flow of gold to the United States created enormous possibilities for credit expansion. At the same time Europe, after the war and during and after the inflation period, was poor in capital, and money rates were extremely high, often because of the risk involved. But the American investor was eager for the high interest rates and the possibility of making a "good investment." And few bothered much about the consequences or the meaning of our large foreign loans.

The development of our foreign trade in the nineteen-twenties therefore was not the result of a normal exchange of commodities and services between nations but to a large extent the consequence of huge international loans. For this reason it rested on insecure foundations. Any disturbance of the flow of capital would of necessity overthrow this unstable equilibrium and permit the cumulating maladjustments to take their course.

It is to be noted, moreover, that the decline of unit labor costs incident to the improvement in technique and the stability of wages served also to intensify the international disequilibrium, enabling, as it did, the United States to sell some commodities widely in other countries, thereby increasing the volume of exports in relation to imports and making our creditor position still more untenable.

Confronted with the crisis thrown up by the war—the disturbed trade relations, the monetary upheaval, and the strained debtor-creditor relations—the United States was suddenly forced to play a leading role in international economic affairs without adequate experience, training, or tradition. Standing between the agrarian and the older industrial countries, we occupied a peculiarly difficult position. As we were export-

ing both agricultural and manufactured products, our economic prosperity was affected by the fortunes both of the agrarian and the industrial nations. We were led to make large loans and investments in both the industrial and the nonindustrial countries. The loans to Europe, while an almost inevitable consequence of the war, proved particularly embarrassing, since the goods which those countries wished to present in payment were in considerable measure competitive with American products. Certain it is that the national policies pursued by the United States after 1920 led to disquieting results in terms of international disequilibria and domestic structural changes, notably those affecting agriculture.

The importance of tariff policy, however, can easily be exaggerated. A candid view of all the facts leads to the conclusion that in the decade of the nineteen-twenties a more enlightened tariff policy could not alone have prevented the collapse. There were too many disturbing factors of overwhelming significance, too many rigidities in the economic structure, other than trade barriers, to permit the necessary adaptation to the changed conditions.

From the standpoint of international economic relations certain dominant facts loomed large in significance. They were (1) the difficulty of the debtor countries in making payment in goods, owing to increasing tariff barriers and other trade restrictions; (2) the drain of \$1,375,000,000 of gold into the United States in the decade 1921-1930; (3) the concealment of the underlying disequilibrium by the overdose of foreign lending, which in turn was stimulated by easy domestic credit conditions (partly due to the effect of the gold inflows on the reserves of the commercial banks, and at times to Federal Reserve Bank policy).

Thus the problems of trade, money, and loans were inextricably interwoven, and with fatal consequences. These consequences may briefly be recounted: (1) the breakdown of the gold standard, first in agrarian debtor countries, and finally in those industrial countries most vulnerable either because of lack of balance in their international accounts or because of the illiquid and insolvent position of the banking structure; and (2) the disastrous fall in world prices consequent upon the damming up of agricultural and other surpluses, the deflationary pressure of depreciated currencies upon gold-standard countries, and the liquidation of unsound banking assets. Here, in brief, are the major causes and consequences of the devastating international disequilibrium which underlay the unprecedented world-wide depression.

The causes of the world-wide collapse, therefore, are numerous and

varied. Among them a restrictive tariff policy did indeed play a disturbing role. A more liberal tariff policy could have minimized the international strain measurably. With lower tariffs and a correspondingly larger volume of internationally traded commodities, the price mechanism could have functioned more effectively. A large volume of world trade relative to the flow of international capital is of great stabilizing value, because with broader channels of trade, strains cannot accumulate so seriously as when the channels of trade are narrow. The larger the volume of world trade, the easier it becomes for the price mechanism to restore a balance when once disturbances, always arising in a highly dynamic world, have developed. No international monetary system can work in a world economy in which measures are taken progressively to shut off imports, while at the same time capital movements are taking place on a scale so huge that any sharp curtailment of foreign lending will cause gaps in the balance of payments of debtor countries of such proportions that they cannot readily be filled by the equilibrating items—gold movements and short-term lending.

Not only would a more liberal tariff policy have eased the strains in the postwar period, but it would also have prevented some of the disrupting factors themselves. The high tariff walls throughout the world stimulated the erection of American plants abroad and thereby contributed to an excess of capital exports. Here indeed the tariff policy of foreign countries was chiefly responsible, but the American tariff policy may indirectly have stimulated high tariffs elsewhere. The postwar policies of the United States and other countries gave Germany, in particular, a popular argument in favor of agricultural protectionism which ran somewhat as follows: "We must develop an excess of exports in order to pay reparations. Our exports are checked by foreign tariffs. Therefore we have no alternative; we must raise our agricultural tariffs, thereby check our food imports, and so produce a favorable balance of trade." Finally other arguments, such as those of national defense and the necessity of protecting the agricultural classes, served to raise European tariffs against farm products. These high tariffs contributed, along with other factors, to dam up huge surpluses in the agricultural countries. Ultimately these surpluses broke through in a calamitous decline in prices of raw materials, and this in turn, along with the cessation of foreign lending, smashed the foreign exchanges in the agricultural debtor countries.

Thus the postwar tariffs profoundly intensified in two ways the disequilibrium which accumulated in the decade of the twenties: (1) they

stimulated the flood of foreign investment; (2) they accentuated the raw-material surpluses.

The Tariff Act of 1930, agitated in 1928, introduced in Congress in 1929 at the height of prosperity, and passed before the depression had become serious, intensified international difficulties. It gave the signal for the collapse of the tariff truce which had been proposed in Europe, and led to tariff retaliations in many countries.

Under the impact of the depression, the fall in prices, and the breakdown of the international gold standard, tariffs almost everywhere were raised and other forms of trade restrictions were adopted, including (1) rigid limitation of imports by means of fixed quotas, and (2) rigorous governmental control over foreign exchange.

THE TRADE AGREEMENTS ACT OF 1934

Confronted with these unprecedented obstacles to our foreign trade, the United States undertook in 1934 the essentially new program of bargaining with foreign countries with respect to tariff and other trade restrictions. Until then, with the relatively unimportant exceptions indicated below, this country had always pursued an autonomous tariff policy, one not determined by negotiation or bargaining, but definitely fixed by Congressional action. Until comparatively recent years, this system of tariff-making placed our foreign trade at no serious disadvantage. In consequence of the large invisible items of interest payments due to foreigners on their investments in this country, together with the large immigrant remittances of the prewar period, we were able to offer a generous supply of dollar exchange with which Europe could buy our products. But in the postwar period we had become a creditor country, and, on balance, heavy interest and dividends were due us from abroad. Moreover, immigrant remittances greatly declined with the virtual cessation of migration to the United States. In the decade of the nineteen-twenties our excess of exports was financed, as we have seen, by a large volume of lending abroad. When this was checked in 1928 and finally came to an abrupt end in 1930, foreign countries were placed under severe pressure to restrict their purchases from the United States. The decline in prices and the currency derangement accentuated this development.

Moreover, structural changes in world trade placed this country in an increasingly less favorable position with respect to her export trade. Formerly our exports, consisting largely of foodstuffs and raw materials,

were in great demand in the highly industrialized countries of Western Europe, and were therefore not vulnerable to serious tariff retaliations. That time has passed. Our agricultural surpluses no longer hold a quasi-monopoly position in the world market. Important competing supplies now come from Argentina, Canada, and Australia. Even American cotton is forced to compete with increasing exports from India, Egypt, and Brazil. In addition, the industrial countries themselves are finding it possible, with the development of artificial fertilizer and scientific agriculture, to become more nearly self-sufficient in food-stuffs. Under these circumstances, the United States is not as free as formerly to adopt whatever tariff policies she pleases against her customers. The current situation calls for international consultation and agreement.

The Trade Agreements Act of June 12, 1934, gave the President wide tariff-making powers extending over a three-year period. Should he find that existing duties or other import restrictions of this or other countries were unduly burdening or restricting the foreign trade of the United States, he could enter into trade agreements with foreign governments. In order to carry out these agreements, he was empowered to proclaim a modification of existing tariff duties, but he could not decrease any duty by more than 50 per cent. The Act provided that the lower duties proclaimed by the President should apply to the imports from all foreign countries, with the proviso, however, that he could withhold such benefits from countries which discriminate against American commerce or which pursue policies tending to defeat the purpose of the Act. The President was empowered to put these agreements into effect without ratification by Congress.

The United States had previously had little experience with the making of reciprocal trade treaties or agreements. Before the enactment of the Trade Agreements Act of 1934, our tariff schedules, with rare exceptions, had been established by Acts of Congress and not by negotiation with foreign governments. Only three reciprocity trade treaties affecting tariff duties had been put into actual effect: one with Canada in 1854; a second with Hawaii in 1875; and a third with Cuba in 1902. Ten additional treaties were negotiated from time to time under the general treaty-making power, but none of these ever became effective. In addition, eleven reciprocity treaties were negotiated under a special provision of the Act of 1897, all of which, however, the Senate declined to ratify. In contrast with this record of failure with treaties requiring ratification by the Senate, twenty-one executive trade agreements, ne-

gotiated under special authorization by Congress and not requiring ratification by the Senate, had been concluded under special provisions of the Tariff Acts of 1890 and of 1897. The first series of these agreements was made largely with the British and Spanish West Indies, the Central American countries, and Brazil, under threat of penalty duties on sugar, molasses, tea, coffee, and hides. Under the terms of the act, no tariff reductions could be made on our side. The second series, negotiated under the Act of 1897, made some reductions in the duties on argols, spirits, wines, and art products from certain European countries. In this case the concessions which could be granted on the American side were limited to the very few products indicated.

Tariff reduction by unilateral action is politically difficult in periods of depression. When millions of persons are unemployed, it is not easy to take action which will have the immediate effect of increasing foreign competition in certain lines. When, however, trade obstacles are reduced both here and abroad simultaneously, as is the case under trade agreements, any temporary loss of employment from increased imports is offset at once by increased employment in the export industries. Tariff bargaining therefore is a safer method of making tariff reductions during a depression. This in part accounts for the policy adopted in the Trade Agreements Act of 1934.

While the American trade agreements program was nurtured in the period of the great depression, the European agreements have frequently been born under favorable economic conditions. The famous Cobden Treaty of 1860 between England and France and the network of commercial treaties to which it gave rise were nurtured in the period of industrial progress, expanding employment, rising prices, and general good times which started in the late forties and culminated in the crisis of 1873. Without this era of rapid expansion, it is doubtful whether the restrictive policy of extreme protectionism could have given way to the era of relatively free trade and liberal commercial policy which the Cobden Treaty ushered into being. These liberating tendencies, in turn, gave buoyancy and power to the rising tide of industrial and business prosperity.

The trend toward freer movements of trade, however, was checked by the prolonged and grueling depressions of the seventies and nineties. In 1881 France returned to the policy of protectionism by enacting an increase in tariff rates which became the basis for a new series of tariff negotiations. In 1892 she established a minimum and maximum schedule of rates, the minimum rates to be granted to foreign countries in return

for concessions, but subject to alterations at any time by the French legislature. In effect, the maximum rates were punitive in character.

The network of trade agreements following the Cobden Treaty had tended to consolidate and stabilize, in a measure, tariff rates among the leading European countries. With the defection of France from the system of tariff consolidation through the fixing of rates by agreement, Germany assumed the leadership in the making of reciprocal trade treaties. Abandoning her former autonomous policy, she turned enthusiastically to the system of tariff bargaining and concluded in 1892 four great treaties with important neighboring countries which were to run for twelve years. These treaties, coming in the relatively prosperous years which preceded the panic of 1893, mark an important recession from the high agricultural rates imposed under the stress of the hard times of the seventies by the Bismarck tariff in 1879. At the termination of these treaties, a series of new trade agreements was negotiated in 1904-1905, which continued in effect until interrupted by the World War.

A major benefit which Europe enjoyed from the system of fixing and binding customs duties by treaty arrangements, first under the leadership of France and later that of Germany, was the relative tariff stability which it afforded. In contrast, we have witnessed, in the postwar period, the utmost liberty—or license—on the part of each nation to erect extreme tariffs and other barriers to trade.

Under the 1934 Act, the United States made (by early 1937) trade agreements with sixteen countries, including in the Western Hemisphere Canada, Cuba, Brazil, Colombia, and various Central American countries, and in Europe France, Sweden, Belgium, Holland, Switzerland, and Finland.

These agreements have made a modest but important beginning toward scaling down the prohibitively high rates of the 1930 Tariff Act. A recent publication of the Tariff Commission, consisting of 44 pages, lists in fine print the individual items upon which the rates have been reduced. While a formidable list covering numerous commodities in all the fifteen tariff schedules, it nevertheless includes only a relatively small proportion of the whole range of dutiable imports.

The Cuban agreement is outstanding among the agreements in the concessions which it gained for American agriculture. Cuba, formerly one of the leading markets for American lard, reduced the duty on this product from 9 cents a pound to 1.45 cents and also gave a substantial reduction on wheat flour. Concessions accruing to the benefit of agriculture, however, are contained in all the agreements. These involve

either tariff reductions, the continuance of duty-free status, or enlarged quotas. These concessions include such items as fruits, dried milk, lard, meat products, vegetables, cotton, and tobacco.

The agreement which offered the largest benefits for American manufactured products was doubtless the Canadian agreement. This was very sweeping in scope, and brought tariff reductions on some 800 American export commodities covering a broad range of manufactured products and also a considerable number of agricultural products. Since this agreement was, in terms of the volume of trade affected, the most important of those made, and since the concessions granted by the United States have particularly been subjected to public discussion, we shall elaborate somewhat upon its provisions.

In the Canadian agreement, quotas were placed upon the amounts of certain competitive products which could be imported into the United States at the lower duty. This applied to such products as cattle, calves, dairy cows, cream, certified seed potatoes, Douglas fir, and Western hemlock. These quotas applied to total imports from all countries. Canada was obliged to take her chances on the total import quota assigned. She was in point of fact, however, in a position to get almost the whole advantage, since these imports come almost wholly from Canada.

In the case of cream, 1,500,000 gallons could be imported annually at the reduced rate of 35 cents a gallon. Any excess importation would pay the old rate of 56.6 cents. Thus the interests of our dairy producers were carefully safeguarded. In the predepression years we imported nearly 4,000,000 gallons and in some years over 5,000,000 gallons. The quota allowed at the new duty was about one seventh of 1 per cent of domestic butter-fat production in the United States. The effect on cream and butter prices therefore could not be significant.

It is to be noted that cream was on the free list before 1921 and that the new duty of 35 cents on a limited importation was still nearly twice as high as that fixed in the Tariff Act of 1922, when the rate was 20 cents a gallon.

It has been argued that if any imports came, no matter how small, our price would fall to the level of the price in Canada. This is wrong. The importation of 1,500,000 gallons is so small that it could scarcely affect the price in our huge market. The exportation of this amount would have a much greater effect on prices in the small Canadian market. The lower American tariff would tend to raise Canadian prices toward the American level. Quotas, small to us, mean much to a country like Canada.

The duty on feeder cattle weighing 700 pounds or more was reduced from 3 cents to 2 cents a pound, the same as the rate in the 1922 Tariff Act. This rate, moreover, was limited to a quota of only three fourths of 1 per cent of the annual American slaughter in 1928-1932. Since Mexican cattle weigh only 400 to 450 pounds, Canada got practically all the benefit of this quota of about 155,000 head. The predepression importation from Canada amounted to from 160,000 to 200,000 head.

A duty of $2\frac{1}{2}$ cents a pound was imposed, in the 1930 Act, on calves weighing less than 175 pounds. Under the agreement this rate was lowered to $1\frac{1}{2}$ cents, the same as the rate fixed under the Act of 1922. Only a limited importation of one fourth of 1 per cent of the annual slaughter of calves in 1928-1932 could come in at this reduced rate. With respect to dairy cows, the agreement provided that 20,000 could be imported annually at a rate of $1\frac{1}{2}$ cents a pound, a fraction of 1 per cent of the United States production. Any excess imports must pay the old rate.

The duty on certified seed potatoes was lowered from 75 cents a bushel to 45 cents for a part of the year and to 60 cents from December 1 to March 1. But only 750,000 bushels could be admitted at this lower rate. This amounted to about 7.5 per cent of domestic production. The duty on potatoes for consumption was not reduced. Farmers generally, who do not grow certified seed potatoes, therefore stood to gain from the agreement.

Lumber was formerly on the free list, even under the Act of 1922. In 1930 a duty of \$1 a thousand feet was imposed on spruce, pine, fir, and hemlock. In the Revenue Act of 1932 an additional excise tax of \$3 a thousand feet was imposed.

Under the agreement these rates are reduced 50 per cent but with the proviso that only 250,000,000 board feet of Douglas fir and Western hemlock may be imported at the reduced rates. This amounts to 5 per cent of our consumption of these species. Other soft woods, spruce and pine, supplement our own inadequate supplies. The competitive imports of Douglas fir and Western hemlock affect chiefly the domestic industry in Oregon and Washington.

Direct benefits to American agricultural producers from the Canadian agreements flowed chiefly to vegetable growers and to citrus fruit-growers. The vegetable concessions, covering a long list of products, benefited the whole tier of Northern states near the Canadian border, and also the Southern states, especially in the off season. In addition Canada made duty reductions on our meat products, Indian corn, oatmeal, timothy seed, potatoes, and sweet potatoes.

While special agricultural groups benefited directly, agriculture as a whole benefited from the increased urban employment in those industries which export to Canada, especially farm machinery, industrial machinery, electrical machinery, and automobiles. All these industries received important duty concessions in the Canadian agreement. The recovery of a major part of the \$600,000,000 of Canadian trade which we had lost since 1930 is of no small significance for our urban prosperity. Farmers cannot achieve normal agricultural prosperity without a substantial increase in urban employment and urban purchasing power. The prices of agricultural products correlate very closely with the rise and fall in urban pay rolls.

An argument frequently made against these trade agreements, and especially the one with Canada, is that it does not make sense to engage in a government program to raise farm prices and then at the same time reduce the tariff on agricultural products. There is in reality no inconsistency, for it is reasonable to suppose that the indirect benefits flowing from the increased urban prosperity to our export industries would benefit agriculture in general. It is always an error to concentrate attention narrowly upon the direct effect of a certain course of action upon a restricted segment of the economy.

The trade agreement concluded between the United States and Canada in November, 1935, marked the end of a period of tariff reprisals which had seriously injured the trade relations of these two neighboring countries. By the Emergency Tariff Act of 1921 and the Tariff Acts of 1922 and 1930, the United States had successively raised the duties against Canadian products. These acts had imposed duties on a large number of commodities which formerly had been imported duty-free from Canada. The rates were made very much higher, many doubled, in the Act of 1930. Canada retaliated by raising her tariff barriers still higher against American products in 1930, 1931, and 1932. The trade on many individual commodities was reduced to insignificant figures, partly in consequence of these tariff increases, and partly for other reasons. Moreover, trade was deliberately shifted from the United States to England. The Canadian tariff differentials against American and in favor of British products were raised materially. Our share in the Canadian market seriously declined after 1930, while England's share materially increased. In addition, the increases in the duties imposed by Canada against our manufactures induced American concerns to establish branches in Canada, thereby reducing the employment of American labor.

THE MOST-FAVORED-NATION POLICY

A word needs to be said about the most-favored-nation policy under which these agreements were negotiated. This policy is much misunderstood. It is said that these bilateral trade agreements are certain to engender international ill will. This would be true if the agreements were made on a preferential basis, resulting thereby in discriminations against other countries. Such, however, is not the case. All countries (except those which discriminate against our trade, such as Germany) are given the same lower duties which we offer the countries with which we have made trade agreements. Equality of treatment is the cornerstone of the most-favored-nation policy. In the case of Cuba, because of special historical relations, we give and receive preferential tariff treatment; but this is an exception. All other trade agreements are based on the principle of equality of treatment to all countries that treat our trade fairly.

But, it may be asked, How is it practical to give the world the same benefits we accord the trade agreement countries? The answer is that the concessions which we have made to any particular country apply mainly to commodities of which that country is the leading supplier. Of the products on which we made tariff reductions in the Canadian agreement, about 95 per cent of the total imports came from Canada. In our other trade agreements, however, the percentage as a rule was considerably lower.

The United States has received many advantages from the fact that these reciprocal agreements are concluded on the most-favored-nation basis. We have the assurance that the concessions other countries have granted us will not be nullified by the granting of still more favorable rates to third countries. This is a point always overlooked by those who superficially conclude that a preferential agreement is more desirable. Several illustrations can be given of the advantages to us of assurance of most-favored-nation treatment by other countries. After we had concluded our agreement with Belgium, that country reduced her duties on certain commodities imported from France. Had we not had a most-favored-nation agreement, this action would have destroyed the value of the concessions made to us on similar commodities. As it was, the United States automatically obtained the benefit of the new low rates accorded by Belgium to France. Similarly, rate reductions made by Belgium to Italy and Germany in the summer of 1935 were automatically extended to the United States. For these concessions we paid noth-

ing, except the assurance that we also would accord most-favored-nation treatment to Belgium. Other benefits have come to the United States in recent months from tariff reduction made to third countries by Sweden, Brazil, Canada, Netherlands, Switzerland, and Honduras, all of which reductions were likewise granted the United States.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The protective tariff is an economic benefit to new countries but an injury to developed countries." Do you agree?

2. "A country should make foreign loans only to countries whose imports are noncompetitive with its own." Discuss the implications of this statement under each of the two following conditions: (a) bilateral balancing of trade and service payments (including interest) between each pair of countries; (b) triangular trade.

3. "In the postwar period, the United States should either have refused to permit foreign lending by its citizens, or else it should have lowered its tariff duties." Argue for or against this statement. What relation is there between foreign lending and tariff policy?

4. "A general lowering of tariff duties throughout the world in the postwar period would have prevented the world depression." Give reasons for agreement or disagreement.

5. "During a period of depression, tariff bargaining is a better device for tariff reduction than unilateral action." Defend this statement.

6. "From the long-run standpoint, a country which reduces its own tariff without securing compensating reductions in the tariff duties of other countries, is merely sacrificing its own interest for the benefit of the rest of the world." With this statement economists generally would disagree. On what grounds?

7. "If the tariff on a certain commodity is removed, the price in the importing country will fall to the level prevailing in the exporting country." Show why this could not be true. Under what conditions might the price fall considerably in the importing country? Under what conditions might the price rise considerably in the exporting country?

8. "Under the most-favored-nation policy, a country is giving away benefits to third countries without getting any benefit in return." Show why this is not true.

9. "Obviously we could obtain more benefits for our exporters if we concluded preferential agreements with foreign countries." Do you agree? What special danger confronts a country which makes preferential agreements?

10. "A program of trade agreements necessarily comes into conflict with the most-favored-nation policy." Do you agree? Why, or why not?

INDEX

- Acceptances, bank, 609
Adamson Act, 496
Administered price, 116, 284
Advertising, 87, 149; as a method of manipulating demand, 196-202
Aftalion, Albert, on income theory of prices, 337-339
Agents of production, distinguished from factors, 60; effect of mobility of, on price, 176-177; marginal use of, 179-181. *See also* Factors of production
Agricultural Adjustment Act, 263, 283, 284
Agricultural prices, 281-285, 520-527
Agriculture, cost of final products of, to consumer, 92, 223; effect of price fluctuations on, 324-325, 520-527; movement of land values in, 520-523; rental value of land in, 520, 522-523; prices in, related to other prices, 521-527; overcapitalization of land value in, 527
American Federation of Labor, fluctuation of membership of, with prices, 328; effect of, on immigration legislation, 480; structure of, 483; policies of, 483-484, 498-499; growth of, 484
Anarchism, 569
Anderson, B. M., Jr., cited, 658
Arbitration, compulsory, in Australia and New Zealand, 498; compulsory, opposition of labor to, 498-499
Authoritarian states compared with communistic, 574
Average product, 63-64
Balance of international payments of the United States for 1935 and 1936, 623
Balance of trade, 619-628; active and passive, 619; of United States as a debtor country, 619-621; favorable and unfavorable, 619-620, 625-627; of United States as a creditor country, 621-625, factors governing, 627-628
Bank deposits, how originated, 311-313; insurance of, 372, 374
Bank drafts, 313
Bank notes, 304-306
Banking, centralization of control of, 372; branch, 373; qualitative versus quantitative control of, 373-375. *See also* Federal Reserve System
Banking Act of 1935, 372, 374
Basing-point system of price-making, 249, 251
Berle, A. A., Jr., on control of corporations, 51
Bernstein, Eduard, on Marxian theory, 568
Bill of exchange, 312; when used in foreign exchange, 608-609
Bimetallism, 306-311; Thomas Amendment, permitting, 310; Silver Purchase Act of 1934 and, 310
Black, J. D., cited, 283
Blanc, Louis, and French socialists, 563
Bland-Allison Act, 308-309
Bloomfield, Daniel, cited, 492
Böhm-Bawerk, Eugen von, criticism of Marxian system by, 567
Bolshevism, 569-570
Bowley, A. L., on distribution of income, 556
Boycott, 493-495
Burns, Arthur R., cited, 250
Business cycles, 348-366; characterized, 348-349; phases of, 349-355; rainfall and, 355-356; self-generating, 356-358; monetary explanation of, 358-359; overproduction and underconsumption theory of, 360-363; investment and, 363-365; an unsolved problem

- lem, 365-366; desirability of elimination of, 366-367; methods of control of, 368-375; in relation to unemployment, 487. *See also* Stabilization
- Business organization, forms of, 45-50
- Business profits, 463-464, 527-532. *See also* Profits
- Capital, accumulation of, 20, 536-537; defined, 55-56; nature and kinds of, 438-441; return to, 441-442; marginal productivity of, 441-443; superior productivity of, 443-445; marginal supply price of, 450-451; cost of, 455; international movement of, 623-624
- Cartel, 242
- Cassel, Gustav, cited, 20, 613
- Causal relations in economics, 9
- Chamberlin, Edward, cited, 186, 200
- Chambers, C. R., cited, 522
- Child-labor laws, 37-38, 499-500
- Circulating media distinguished from money, 302
- Clark, Colin, cited, 552, 557
- Clark, F. E., on marketing functions, 86-87
- Clayton Act, 255, 256, 257
- Closed shop, 481, 490-491
- Coinage, 303-304, 306-310
- Collective bargaining, 489-491
- Commercial policy of the United States, *see* International trade, Tariff, and Trade agreements
- Committee on Economic Security, 502
- Committee for Industrial Organization, 484
- Commodity Exchange Act, 295
- Communism, in Russia, 569-573; contrasted with socialism, 573-574
- Communist Manifesto*, 564-567
- Company union, 490
- Comparative cost, law of, 629-643; in regional trade, 630-633; in international trade, 633-639; summary of theory of, 639
- Competition, perfect and imperfect, contrasted, 112; contrasted with monopoly, 186-189; monopolistic, 193-196; and manipulation of demand, 196-202; imperfect, and wages, 398-401; natural limitations on, 473-474; foreign, 640-643
- Competitive price, 94-185; the dynamics of, 203-240. *See also* Normal price
- Concentration in industry, *see* Large-scale production
- Concentration of wealth and income, *see* Inequalities of property-ownership and income
- Consolidation, 243
- Conspiracy and labor disputes, 493-495
- Constant cost, defined, 213; conditions necessary for, 214-217; shifting of costs and, 217-218
- Consumers' demand and utility, 134-150
- Consumption, compared with saving, 18; distribution of, compared with income, 539
- Co-operation, 47-48
- Corporations, organization of, 46-47; and control of economic life, 50-52
- Cost, lowest unit, 65-71; of production, meaning of, 152-156; of production and price, 152-167; prime and supplementary, 160-167; overhead, distinguished from joint, 182-183; constant, 213-218; varying, under dynamic conditions, 213-238; increasing, 218-226; decreasing, 227-238
- Credit instruments, 312-313
- Crises, *see* Business cycles
- Currency, *see* Money
- Customs duties, *see* Tariff
- Damage suits in labor disputes, 494-495
- Davis, J. S., cited, 283, 284
- Day, E. E., cited, 515
- Debt settlement, *see* International debts
- Debt of United States government, 595; of state and local governments, 595
- Decreasing cost, defined, 227; causes of, 228-233; distinguished from declining cost, 233-234; limitations on, 234-235; and monopoly, 235-238
- Demand, defined, 98; individual, 98; collective, 102-103; elastic and inelastic, 104-108; changes in, 109-110, 121-122, 125; analysis of, 134-136; relation of marginal utility to, 143-148; effect of changes in, on earnings of fixed capital, 162-167; manipulation of, through advertising, 196-202; effect of changes in, on cost, 213
- Demand price, 97-99

- Demand schedules, 99-100, 103-108
- Demand and supply, law of, *see* Supply and demand
- Deposit currency, 311-314
- Differential costs, entrepreneurs', 204-209; and imperfect competition, 209-210; improvements and, 211-212
- Diminishing returns, law of, 57; and constant cost, 214; relation of, to increasing cost, 218; in relation to rent, 424-425; from the use of capital, 453-454
- Diminishing utility, principle of, 139-141; relation of, to progressive taxation, 502
- Distribution of estates, in Massachusetts, 540-546; in United States, 541-545; in Minnesota, 546-547
- Distribution of income, defined, 379-380; functional, 379-475, 558-559; importance of, 380-381. *See also* Interest, Profits, Rent, and Wages
- Distribution of income (personal), 380-381, 539-540; dynamic aspects of, 510-537; in United States, 548-551, 556-557; equalitarian, 551-552; relative permanency of, 553-557; causes of inequality in, 556-557; in Great Britain, 554-557; personal contrasted with functional, 560
- Distribution of wealth among living persons, 547-551. *See also* Distribution of estates
- Division of labor, 23-26; and international trade, 629
- Domestic system, 477-478
- Economic law, 9-12
- Economic order, the present, 27-29; waste in the present, 576-577
- Economics, as a science, 5-12; subject matter of, 12-15
- Efficiency versus unfair practices in relation to trusts, 246-247
- Elasticity of the currency, 316-317
- Elasticity of demand, defined, 104; causes of varying, 106-108; effect of, on price and gross receipts, 121-122; and effect of improvements on wages, 403; for labor, 482; relation of, to shifting of taxes, 586
- Emergency Transportation Act of 1933, 280
- Employers' associations, 486
- Engels, F., 564-567
- Entrepreneur, function of, 28, 42-44, 464-467; marginal, 65-66; superior, 66-71; as a risk-taker, 470-472
- Entrepreneurs' differential costs, 204-212; and increasing cost, 221-222
- Equation of exchange, 332-335
- Erdman Act, 496
- Evolutionary socialism, 567-568
- Exchange, foreign, rates of, 611-615
- Exchanges, commodity and security, regulation of, 294-297
- Expenses of production, and price, 152-167, 169-185, 213-238. *See also* Cost
- Exploitation, gains of, 474-475
- Exports, American, increase of, during World War, 621; composition of, 647-650
- Fabian socialism, 567
- Factors of production, classification of, 55-56, scarcity of, 56-57; combination of, 57-60; cost of, 155-156; combination of, and decreasing cost, 228-230; and entrepreneurship, 464-466. *See also* Agents of production
- Factory legislation, 37-40
- Family-allowance wage, 480-481
- Farm prices, *see* Agricultural prices
- Federal Deposit Insurance Corporation, 372, 374
- Federal reserve bank notes, 305-306
- Federal reserve notes, 315-316
- Federal Reserve System, 314-317; effect of, on monetary system, 342-346; relation of, to business cycle control, 370-372
- Federal Trade Commission, 255
- Fees, 582
- Filene, E. A., on minimum wage, 482
- Fisher, Irving, on business cycles, 359; on time preference, 446
- Five-year plans in Russia, 571-572
- Foreign exchange, 607-617; rate of, defined, 607-608; forms of, 608; practices, 608-611; rate of, how determined, 611-615; pegged rate of, 615-616; government control of, 616-617; International Gold Agreement of 1936 and rates of, 617
- Foreign trade, *see* International trade

- Fourier, Charles, 563-564
 Free-trade principles, 644-659. *See also* Tariff
 Friday, David, on income of corporations, 530
 Functional distribution, *see* Distribution
 Functions of money, 301
 Futures, dealings in, 291-292

 Gide, Charles, on law of supply and demand, 35
 Gold, as a standard of value, 302-303, 306-311; noncirculation of, in United States, 303-304; international movements of, since 1914, 343-344
 Gold certificates, 304
 Gold parity, 611-612
 Gold Reserve Act of 1934, 302-304, 310
 Gold-shipping points, 611-612
 Gold Stabilizing Fund, 371-372
 Gold Standard Act of 1900, 309
 Goods, durable and nondurable, 19-21; classified as luxuries and necessities, 107; perishable, 111, 113; utility of, 136-138; free and economic, 138
 Government debt, 595-596
 Government finance, economic principles applicable to, 579-580
 Government intervention in labor disputes, 495-499
 Government receipts and expenditures, 591-594
 Government revenue, nontax, 581-582; tariff as a source of, 644. *See also* Government receipts and expenditures
 Governmental enterprises, 48-49
 Grain Futures Act, 294
 Greenbacks, 305, 308
 Gresham's law, 307, 308
 Guild socialism, 569
 Gulick, C. A., Jr., cited, 248, 251

 Harvard Economic Society, cited, 532
 Hawtrey, R. G., on business cycles, 358
 Health insurance, 500-501
 Hedging, 293-294
 High-wages argument for protective tariff, 652, 655-657
 Hobson, John A., on business cycles, 360-363; on minimum wage, 482-483
 Holding company, 243

 Home-market argument for protective tariff, 651-652, 653-655
 Hours of labor, and shorter work-day, 487; legislation concerning, 500
 Human behavior in economic affairs, 8-9

 Immigration, effect of, on wages, 518-519
 Imperfect competition, 112, 116, 128, 209-210, 398-401
 Imports into the United States, ratio of duties on, to total dutiable, 645-647; importance of, 647-650; character of, in 1935, 647-650. *See also* Tariff
 Improvements, and differential costs, 211-212; in relation to decreasing cost, 232-233; and wages, 401-405; on land, return from, 421
 Incidence of taxes, 583-587
 Income, family utilization of, 17; capitalization of, 455-456; earned and unearned, 598; distribution of, *see* Distribution of income
 Income tax, proportion of personal income taken by, in state of New York, 552; Federal, individual, 601
 Increasing cost, 218-226; defined, 218; conditions giving rise to, 219-221; and entrepreneurs' differential costs, 221-222; forces counteracting, 222-224; and incidence of new unit costs, 224-226
 Index numbers, 319-323; purposes of, 319; how constructed, 321-322; of wholesale prices, in United States and England, 322-323
 Individual proprietorship, 45
 Individualistic order, 34
 Individualistic philosophy, 35-36
 Industrial combination and monopoly, 240-260
 Industrial Disputes Act, Canadian, 498
 Industrial fluctuations, *see* Business cycles
 Industrial regimentation, alternative to monetary control, 368
 Industrial Revolution, 37-38, 392, 477-478; effect of, on level of well-being, 553
 Inequalities of property-ownership and income, 539-560
 Infant-industry argument for protective tariff, 650-651, 652

- Injunction in labor disputes, 495
 Installment selling, 336, 663
 Institutional economics, 54-55
 Institutions in the present order, 34
 Interdependent prices, 169-183
 Interest, 438-461; defined, 438; time preference and, 445-447, 450; productivity of capital and, 450; supply of savings and, 450-451; determination of, 451-455, 461; types of, 457; implicit and explicit, 457; loan market and, 458-459; rate of, 459-461; and the supply of money, 459-461; under dynamic conditions, 532-537
 International debts, 617, 620-625, 602-664
 International Gold Agreement of 1930, 617
 International trade, and division of labor, 629; basic reason for, 629-630; causes of, 632-633; determined by relative productivity of labor, 634-639; fear of foreign competition in, 639-643; importance of, for United States, 647-650. *See also* Balance of trade, Comparative cost, Foreign exchange, Tariff, and Trade agreements
 Interstate Commerce Commission, 277, 279
 Inventions, effect of, on labor productivity, 392; combating tendency to diminishing returns, 431. *See also* Improvements
 Investing class, 532-537
 Investment, effect of, on business cycles, 363-365
 Jaurès, Jean, 568
 Jevons, W. S., on business cycles, 355
 Johnson, Judge A. W., cited, 546
 Joint Commission of Agricultural Inquiry, on cost of marketing, 25, 90-91
 Joint cost, 181-183; distinguished from overhead cost, 182-183
 Joint demand, 170; price and, 175-179
 Joint supply, 171; price and, 181-182
 Kansas Industrial Court, 497-498
 Keynes, J. M., on business cycles, 363
 King, W. I., cited, 91, 544, 557, 663; on distribution of wealth, 547; on proportion of national income saved in United States, 552; on functional shares in distribution, 558-559
 Knights of Labor, effect of, on labor legislation, 480; history of, 483
 Kuczynski, R. R., cited, 395
 Labor, demand for, determined by marginal productivity of, 383-388; competition of machinery with, 388-389; supply of, 390-393; wages of, affected by improvements, 401-405; marginal productivity of, affected by trade unionism, 470-483; marginal productivity of, affected by legislation, 506-507
 Labor disputes, government intervention in, 495-499. *See also* Strikes
 Labor-income groups, 407-410
 Labor legislation, 495-508; minimum-wage laws, 499-500; hours of work, 500; social insurance, 500-505; economic principles of, 506-508. *See also* Social Security Act
 Labor problems, *see* Collective bargaining, Labor legislation, Social Security Act, and Trade unionism
 Labor-saving machinery, effect of, on wages, 401-405; effect of, on employment, 488-489
 Laidler, H. W., cited, 571
 Laissez faire, 34-37, 39
 Land, use of term in economics, 420; different functions of, 420; demand for, 423-425; marginal productivity of, 425-426, 430; rent and differences in, 427-430; value of, as affected by dynamic conditions, 520-527
 Large-scale production, 71-76; and decreasing cost, 234-235; as a factor in monopoly, 235-238
 Lassalle, Ferdinand, 563
 "Lawful money," 306
 Leven, Maurice, on utilization of family incomes, 17
 Lewis, John L., 484
 Liquidity of bank assets, 373-375
 Loan market, 458-459
 Lowest unit cost, 65-71
 Machine methods and increased productivity of labor, 443-445

- Malthus, T. R., on free trade, 12; on population, 390-393
- Marginal buyers and sellers, 126-128
- Marginal cost, 68-70, 158-159. *See also* Prime and supplementary cost
- Marginal entrepreneur, 65-66, 206-209
- Marginal product of variable agents, 63-64; in relation to price, 179-180
- Marginal productivity, of labor, 383-388; of land, 425-426, 430; of capital, 441-443, 456; in relation to international trade, 633-639
- Marginal revenue curve, 131
- Marginal use of agents of production, in relation to price, 179-180
- Marginal utility, defined, 142; relation of, to individual demand, 143-148; market price and, 148-149; practical applications of the theory of, 149-150; distribution of the burden of taxation and, 150, 601-602
- Margins, extensive and intensive, 428
- Market, defined, 81; various concepts of, 81-83; producers' and consumers', contrasted, 83-86; sharing the, a method of control, 252-253
- Market price, 118-132; defined, 96; contrasted with normal price, 97; without reservation of supply, 118-123; with reservation of supply, 124-125; effect of changes in demand and supply on, 125-126; equilibrium, and actual, contrasted, 128; marginal utility and, 148-149
- Marketing, 81-93; functions of, 86-89; specialization and, 89-93; retail, cost of, 90-93
- Markets, related, 85
- Marshall, Alfred, on utility, 140; on activities and wants, 149; on successful independent manager, 465; on economic analysis, 473
- Marx, Karl, 563, 564-567
- Materialist interpretation of history, 564-565; criticized, 566-567
- Means, G. C., on control of corporations, 51
- Mediation and conciliation, 496-497
- Mergers in industry, 243
- Mill, John Stuart, 13; on law of diminishing returns, 57; on quantity theory of money, 335
- Minimum-wage legislation, 418-419, 499-500, 507-508
- Mitchell, W. C., on business cycles, 356-358; cited, 552
- Monetary history of United States, 306-311
- Monetary purchasing power, control of flow of, 368-375
- Monetary standards, 302-303
- Monetary system of United States, 317
- Monetary theory, 330-347
- Money, functions of, 301; defined, 301-302; and circulating media, 301-302; standards of, 302-303; kinds of, in United States, 303-306; history of, in United States, 306-311; deposit currency as, 311-314; theory of value of, 330-332; quantity theory of, 335-337; income theory of, 337-339; supply of, and interest rates, 459-461
- Monopolistic competition, 193-196, 248-253, 260
- Monopoly, defined, 186; equilibrium price and, 186-202; causes of, 189-190; and price discrimination, 190-193; decreasing cost and, 235-238; industrial combination and, 240-260; tax on revenue from, 585
- Monopoly price, 128-131, 186-202, 235-238, 247-254; regulated, 265-281
- Moore, Henry L., on business cycles, 355
- Most-favored-nation tariff policy, 674-675
- Moulton, Harold G., on utilization of family incomes, 17
- National bank notes, 305
- National Bureau of Economic Research, on national income, 16; cited, 19; on distribution of income, 549, 550, 559
- National-defense argument for protective tariff, 652-653
- National income, amount of, 16; division of, between durable and nondurable goods, 19; in relation to taxation, 593; distribution of, *see* Distribution of income
- National Industrial Conference Board, cited, 512; on cost of government, 590-591, 593
- National Industrial Recovery Act, 256-257, 260, 263, 282, 368, 499; regulation of hours of labor by, 500
- National Labor Union, 483

- National Trades' Union, 483
 Newlands Act, 496
 New Zealand, arbitration of labor disputes in, 498
 Nontax revenue, 581-582
 Normal price, defined, 97; and cost of production, 152-167; as equilibrium price, 153-154; relation of, to prime and supplementary cost, 164-165. *See also* Price
 Normal value, summary of theory of, 184-185. *See also* Normal price and Price
 Norris-La Guardia Act, 495
 Nourse, E. G., cited, 283
- Old-age insurance, 501, 502-503
 Old-age pension, provision of Social Security Act for, 501
 Open-shop controversy, 490-491
 Opportunity cost defined, 155
 Organization of production, 42-52
 Overhead cost distinguished from joint cost, 182-183
 Overproduction theory of business cycles, 360-363
 Owen, Robert, 563-564
- Par of exchange, 607
 Partial monopoly, 199-202. *See also* Monopoly
 Partnership, 45
 Patents as bases for monopoly, 247-248
 Persons, W. M., cited, 515
 Picketing, 494
 Pigou, A. C., on proportion of national income saved in United Kingdom, 552; cited, 557
 Political units as entrepreneurs, 48-49
 Pools, 242-243
 Population, Malthusian theory of, 390-391; and the supply of labor, 390-393; growth of, in England, France, and United States, 391-393; decline in growth of, 394-395
 Price, defined, 94; analysis of, 95-96; market and normal, contrasted, 97; equilibrium, under monopoly, 128-131, 186-190; relation of, to cost of production, 152-167; interdependent prices, 169-183; competitive, the dynamics of, 203-212; and speculation, 287-297. *See also* Cost, Monopoly, Normal price, and Value
 Price fluctuations, 319-328; effect of, on borrowers and lenders, 324-325; effect of, on investors, 325-326; effect of, on employers and wage-earners, 326-328; in relation to business cycles, 359; and wages, 513-517
 Price leadership, 248-252
 Price movements, theories of, 339-341; since 1914, 341-347
 Prices, income theory of, 337-339; collapse of, after 1929, 344-345; future of, 346-347; public, 581
 Prime and supplementary cost, 160-167
 Production, guided by consumption, 18; defined, 21; regulation of, 34-40; organization of, 42-52; for the future, 287-288; superiority of the roundabout method in, 443-445
 Productive activities, 22-26
 Productivity, effect of, on wages, 512-513. *See also* Marginal productivity
 Profits, 463-475; "pure," distinguished from "business," 463-464; as a share in distribution, 464-467; from earnings of rare natural ability, 467-468; relation of risk to, 468-472; real cost of risk-taking and, 472-473; from imperfect competition, 473; from monopoly and exploitation, 474-475; history of, in United States, 527-532
 Property, kinds of, 31; as a social institution, 557-558
 Proprietorship, individual, 45
 Public regulation, 36-37, 262-285; growing importance of, 262-263
 Public utilities, 262-281; nature of, 263-265; regulation of rates of, 265-273; market value of property of, as basis for valuation of, 266-267; reproduction cost as basis for valuation of, 267-270; rate of return on property of, 270-271
 Public Utility Act of 1935, 14
 Purchasing power, control of flow of, by budgetary policy, 368-370; control of flow of, through Federal Reserve System, 370-375
 Purchasing-power parity, 612-614
- Quantity theory of money, 335-337

- Railroad rates, 267, 277-281
 Railway Labor Disputes Act, 496-497
 Rainfall and business cycles, 355-356
 Rate of return on property of public utilities, 270-271
 Rates of public utilities, regulation of, 265-273
 Reciprocal trade agreements, 668-670
 Reconstruction Finance Corporation, 282
 Rent of land, 420-436; defined, 420; explicit and implicit, 421-422; method of determining, 422; Ricardian theory of, 422-423; and demand for land, 423-425; relation of, to diminishing returns, 424, 428, 429-431; determined by marginal productivity, 425-426; and differences in land, 427-430; generalized theory of, 430-431; agricultural, 432; urban, 433-434; considered an unearned income, 434-436; effect of dynamic conditions on, 520-527; tax on, 584
 Reparation payments, effect of, on foreign exchange, 613-614
 Reproduction cost, 267-270. *See also* Cost, Normal value, and Price
 Resale price maintenance, 258
 Reserve requirements of Federal Reserve System, 314-315
 Ricardo, David, on free trade, 12; on importance of distribution, 380; on determination of rent, 423
 Risk, by whom assumed in business, 43; and entrepreneurs' differential costs, 207; and production for the future, 287-288; speculation and, 290-292; assumption of, by entrepreneur, 470; profits and, 471-472; real cost of, 472-473
 Rival demand, defined, 170; and price, 179-181
 Rival supply, defined, 170; and price, 171-175
 Robertson, D. H., on business cycles, 363
 Robinson, Joan, cited, 186
 Robinson-Patman Act, 257
 Rowe, H. B., cited, 284
 Rühl, Alfred, on composition of imports into United States, 650
 "Rule of reason," 244-245
 Russia, communism in, 569-573; land nationalization in, 570-571
 Saving, and consumption contrasted, 18; and production, 28-29; motives for, 447-450; involuntary, 448-449; by corporations, 449; factors affecting, 449; supply of, 450-451; marginal cost of, 450-451; demand for, 452-455; social, 536-537
 Scarcity, a condition to value, 138-139
 Scrip marks, 617
 Seager, H. R., cited, 248, 251
 Secrist, Horace, on related markets, 85
 Securities Exchange Act, 296-297
 Self-interest, a factor in social control, 40
 Sherman Antitrust Law, 243-245, 255, 256
 Sherman Silver Purchase Act, 305, 309
 Sidgwick, Henry, on the tariff, 653
 Silver Purchase Act of 1934, 310
 Single tax, 435-436
 Smith, Adam, on laissez faire, 34; on unregulated private capitalism, 37; on tariff, 653
 Social control, 36-40
 Social income, 31-32
 Social insurance, 500-501
 Social legislation, in England, 37-39; in United States, 39-40
 Social Security Act, 501-505
 Social utility and property institutions, 557-558
 Socialism, 562-578; extent of, 562; origin and development of, 563; utopian, 563-564; Marxian, 564-567; Fabian, 567; evolutionary, 567-568; guild, 569; contrasted with communism, 573-574; contrasted with capitalism, 574; criticized, 574-576; value and distribution under, 577-578. *See also* Communism
 Special assessments, 581, 582
 Specialization and marketing, 89-93
 Specie parity, 607
 Speculation, and prices, 287-297; and risk-taking, 290-292; effect of, on price, 292-294; in foreign exchange, 614-615
 Spiethoff, Arthur, on business cycles, 363
 Stabilization, 368-375. *See also* Business cycles
 Stamp, Sir Josiah, on proportion of national income saved in United Kingdom, 552; on distribution of income in Great Britain, 553-556

- Standard of value, 306-311
- Static state, competition in, and the individual producer, 203-204
- Stock market, speculation on, 290-292; regulation of, 294-297; prices on, 528-530
- Strikes, 489
- Style goods, 113, 201
- Subsidy, effect of, under increasing cost, 225-226; to agricultural products, 226
- Subsistence theory of wages, 395-396
- Substitution, conditions determining possibility of, 172; effect of, on price, 172, potential, 173; complete and partial, 173-174; relation of, to cost of production, 174-175; principle of, in distribution, 388-389, 422
- Supplementary cost, *see* Prime and supplementary cost
- Supply, defined, 111; schedules of, 111-115; determination of, 115-116, 152; changes in, 116-117; unreserved, 118-123; reserved, 122-125; effect on market price of changes in, 126
- Supply and demand, law of, 34-35, 97, 125; as regulative forces, 34-35; interaction of, 125-126
- Supply price, defined, 111; significance of, 123-124
- Syndicalism, 568-569
- Taft, W. H., on collective bargaining, 491-492
- Tariff, principles of, 644-660; history of, in United States, 644-647; percentage of imports into United States free from, 647; arguments for protective, 650-652; arguments for protective, criticized, 652-657; on agricultural commodities, 654; industries injured by, 657-658; British and American policies of, compared, 659-660; post-war policy of United States regarding, 662-667; reciprocity treaties of United States regarding, 668-669; most-favored-nation policy of, 674-675
- Tax, defined, 581; kinds of, 582-583; direct and indirect, 582-583; shifting and incidence of, 583-587; on rent, 584; on surpluses, 584; on net income, 584-585; on margins, 585; on monopoly revenue, 585; capitalization of, 587-588; amount of, 588-590; on net income of individuals (Act of 1936), 601
- Taxation, relation of, to distribution of wealth and income, 580; determination of amount of, 588-590; increase in, in United States, 590-595; tests of a good system of, 596-597; distribution of burdens of, 597-600; benefit theory of, 597-598; ability-to-pay theory of, 598-599; equal-sacrifice theory of, 599; progressive versus proportional, 600-604
- Thomas Amendment, 310
- Time preference, 445-447; defined, 446; individual differences in respect to, 446, ratios of, 446-447; relation of, to rate of interest, 454-455
- Tourist marks, 617
- Trade agreements, of United States, 662-675; under Act of 1934, 667-673; early, 668-669; European, 669-670; effect of, on United States trade, 671-673
- Trade associations, 253-254
- Trade unionism, 477-499; origin of, 477-478; effect of, on wages, 479-483, 520; growth of, 483-484; extent of organization of, 485; strongholds of, 485-486; demands of, 486-489; methods of, 489-491; attitude of employers in United States and Great Britain toward, 491; place of, in modern industrial order, 491-493; legal status of, 493-495
- Transportation Act of 1920, 279-280
- Transportation Act of 1933 (Emergency Transportation Act), 280
- Trust policy in United States, 254-259
- Trusts (industrial combinations), 242-243
- Uncertainty, and production for the future, 287-288; and profit, 468-470. *See also* Risk
- Underconsumption theory of business cycles, 360-363
- Unearned increment, 434-436
- Unemployment, causes of, 487-489; extent of, 488-489; insurance against, 500-501, 504-505; machinery and, 567
- Unfair practices as sources of control, 246-247

- Unionism, *see* Trade unionism
- Utility, economic, 26; control over, 29-31; consumer demand and, 134-150; defined, 136; economic, explained, 136-138; principle of diminishing, 139-141; marginal, 141-143; effective, distinguished from total, 142
- Utopian socialism, 563-564
- Valuation of public utilities, *see* Public utilities
- Value, 81-297; of money, 330-332. *See also* Price
- Victoria (Australia), minimum-wage boards in, 498
- Wage adjustments, principles of, 405-406
- Wage differences, 407-419; apparent, 409-410; proximate causes of, 410-411; fundamental causes of, 411-413; between different places, 413-417; effect of mobility on, 416
- Wage rates, in building trades, 414; in cotton textiles, 415
- Wages, 379-419; defined, 381; for personal services, how determined, 382; different from other prices, 382-383; determined by marginal productivity, 383-388; and the principle of substitution, 388-389; subsistence theory of, 395-396; and imperfect competition, 398-401; effect of trade-unions on, 399, 400, 479-483, 520; effect of improvements on, 401-405; low, not always cause of high profits, 416-417; of women, 417-418; long-run trend of, 510-511; effect of increased productivity on, 512-513; effect of price fluctuations on, 513-517; effect of immigration on, 518-520
- Wagner Labor Relations Act, 497
- Wants, and production, 21-23; expansion of, 149
- War Labor Board, 497
- War loans, *see* International debts
- Warburton, Clark, on utilization of family incomes, 17
- Watson-Parker Act, 496
- Wealth, distribution of, 379, 540-547. *See also* Distribution of income
- Wicksell, Knut, on business cycles, 363
- Wieser, F. von, cited, 337; on entrepreneurship, 465
- Wolman, Leo, on union membership, 484
- Workmen's compensation laws, 40, 500-501
- Wright, Judge J. W., cited, 546

