लाल बहादुर शास्त्री प्रशासन अकादमी
Lal Bahadur Shastri Academy
of Administration
मसूरी
MUSSOORIE
पुस्तकालय
LIBRARY
अवाप्ति संख्या 105563
Accession No. [redacted]9
वर्ग सख्या
Class No. 330.1
पुस्तक संख्या
Book No. Tau 4th ed.

PRINCIPLES OF ECONOMICS

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO
MACMILLAN AND CO., LIMITED
LONDON · BOMBAY · CALCUTTA · MADRAS
MELBOURNE
THE MACMILLAN COMPANY
OF CANADA, LIMITED
TORONTO

PRINCIPLES OF ECONOMICS

by F. W. TAUSSIG
LL.B., PH.D., LITT.D., LL.D.
Sometime Henry Lee Professor of Economics in Harvard University

VOLUME I
FOURTH EDITION

New York
THE MACMILLAN COMPANY
1945

FOURTH EDITION COPYRIGHTED, 1939,
BY THE MACMILLAN COMPANY

All rights reserved - no part of this book may be reproduced in any form without permission in writing from the publisher, except by a reviewer who wishes to quote brief passages in connection with a review written for inclusion in magazine or newspaper.

PRINTED IN THE UNITED STATES OF AMERICA

Published May, 1939. Reprinted
December, 1944; March, August, 1945.

Previous editions copyrighted 1911, 1915, and 1921
By The Macmillan Company

Renewal copyright, 1939, by F. W. Taussig
Renewal copyright, 1943, by William G. Taussig

PATRI · DILECTO
FILIVS · GRATVS

PREFACE TO THE FOURTH EDITION

For the present edition, this book has been revised thruout. The greatest changes are in the two groups of subjects which occupy the most space,—Money and Banking (Book III) and Distribution (Book V). In both of these, the order of presentation is now different. In both all the matter has been revised, new chapters have been added, and others have been dropped. Elsewhere the changes are less radical, and occasionally a chapter has been left almost as it was in the earlier editions; but none has been left quite untouched.

With all the overhauling, the book remains in essentials such as it was when written thirty years ago. In view of the enormous economic and social changes of the period since 1914, the treatment of hardly any subject could be quite the same. But the book is still addressed to the same class of readers as at the outset,—intelligent and educated persons who wish to begin on a systematic study of the subject. It might have been wiser at the outset if I had chosen as its title "Introduction to Economics." But the more pretentious "Principles of Economics" has a long sanction of usage, and convenience of reference is served by its retention.

The question that has caused me most trouble is the extent to which there should be discussion of the developments in theory that have come during the last twenty-odd years. These naturally are of special interest to the professional economists, and the failure to take cognizance of them may seem to them a sad lack. Yet in the main I have not undertaken to discuss them. The chief reason for this restriction is that the new speculations and reasonings have not yet been brought to a consensus of opinion. Anything laid down on them now may well be obsolete within a few years. Economic science is still in the making and remaking. The study of the various current movements of thought is best reserved to a later stage in the reader's grappling with the subject at large. It is as true of the structure of economic theory which is now being

reared as it is of the economic system described in this book that much has to be unlearned and much has still to be learned.

My indebtedness is great to many friends and associates,—so many that an enumeration of them would be tedious. I have to express more specifically my thanks for the aid I have had from some who have not only criticized and suggested but have contributed drafts on one subject or another. Dr. John M. Cassels made the first draft of the chapter on the Individual Firm, and I have made no changes of importance in what he did. To Professor Karl Anderson chiefly belongs the credit for the radical rearrangement of the subject matter of Book III (Money and Banking); and he undertook the first redraft of many passages in that Book. Mr. E. P. Hutchinson has aided me in bringing to date the statistical matter and the text of Chapters 54 and 55, on Population; and Mr. J. S. Bain has given similar aid for Chapter 56, on Inequality. Dr. John P. Miller has been of similar service, especially in Book II; and his aid has been invaluable in the proofreading of the whole.

F. W. TAUSSIG

CAMBRIDGE, MASS.
March, 1939

PREFACE TO THE FIRST EDITION

I HAVE tried in this book to state the principles of economics in such form that they shall be comprehensible to an educated and intelligent person who has not before made any systematic study of the subject. Tho designed in this sense for beginners, the book does not gloss over difficulties or avoid severe reasoning. No one can understand economic phenomena or prepare himself to deal with economic problems who is unwilling to follow trains of reasoning which call for sustained attention. I have done my best to be clear, and to state with care the grounds on which my conclusions rest, as well as the conclusions themselves, but have made no vain pretense of simplifying all things.

The order of the topics has been determined more by convenience for exposition than by any strict regard for system. In general, a subject has been entered on only when the main conclusions relating to it could be followed to the end. Yet so close is the connection between the different parts of economics that it has been necessary sometimes to go part way in the consideration of matters on which the final word had to be reserved for a later stage. Taxation has offered, as regards its place in the arrangement, perhaps the greatest difficulties. It is so closely connected with economics that some consideration of it seemed essential; whereas public finance in the stricter sense, whose problems are political quite as much as economic, has been omitted. Yet a suitable place for taxation was not easy to find. I concluded finally to put the chapters on this subject at the very close, even though they may have the effect of an anti-climax coming as they do after those on socialism.

The book deals chiefly with the industrial conditions of modern countries, and most of all with those of the United States. Economic history and economic development are not considered in any set chapters, being touched only as they happen to illustrate one or another of the problems of contemporary society. Some topics to which economists give much attention in discussion

among themselves receive scant attention or none at all. I have omitted entirely the usual chapters or sections on definition, methodology, and history of dogma; and have said little on such a topic as the subjective theory of value, which in my judgment is of less service for explaining the phenomena of the real world than is supposed by its votaries. These matters and others of the same sort are best left to the professional literature of the subject. I hope this book is not undeserving the attention of specialists; but it is meant to be read by others than specialists.

Tho not written on the usual model of textbooks, and not planned primarily to meet the needs of teachers and students, the book will prove of service, I hope, in institutions which offer substantial courses in economics. The fact that it is addressed to mature persons, not to the immature, should be an argument in favor of such use rather than against it. Being neither an encyclopedic treatise nor a textbook of the familiar sort, it offers no voluminous footnotes and no detailed directions for collateral reading. When facts and figures not of common knowledge have been cited, my sources of information have been stated. At the close of each of the eight Books into which the whole is divided, I have given suggestions for further reading and study.

I have expressed in the text, as occasion arose, my obligations to the contemporary thinkers from whom I have derived most stimulus. For great aid in revising the manuscript and proof, on matters both of form and substance, I am indebted to my colleagues Drs. R. F. Foerster and E. E. Day of Harvard University.

F. W. TAUSSIG

HARVARD UNIVERSITY
March, 1911

CONTENTS

BOOK I

THE ORGANIZATION OF PRODUCTION

CHAPTER PAGE

1. WEALTH AND LABOR 3

§ 1. The subject matter of economics.—§ 2. Wealth; free goods; economic goods. Wealth and welfare.—§ 3. Goods may become economic thru mere scarcity; but commonly do so because of the exertion of labor.—§ 4. Activity may be irksome or pleasurable. Labor usually is continuous, little varied, irksome.—§ 5. Some sorts of labor always pleasurable Economic welfare; happiness.—§ 6. The irksomeness of most labor to be lessened by better general opinion, and by greater leisure thru shorter hours.

2. OF LABOR IN PRODUCTION 15

§ 1. Labor given to material objects deemed alone productive by earlier English economists. Objections to this view.—§ 2. Labor creates utilities only; all labor that issues in utility is productive. Is there nonmaterial wealth?—§ 3. Is there any unproductive labor? Labor given to things harmful.—§ 4. Labor of judges and lawyers; of soldiers.—§ 5. Predatory labor. "Business." The law and unproductive labor.

3. THE DIVISION OF LABOR AND THE DEVELOPMENT OF MODERN INDUSTRY 29

§ 1. Two forms of the division of labor: the simpler and the more complex.—§ 2. Advantages from the simpler form: dexterity, continuity, adaptation to aptitudes.—§ 3. Advantage from the more complex form: the development of machinery. The industrial revolution of the eighteenth century. The use of nature's power.—§ 4. Division of labor means unconscious coöperation. Exchange.—§ 5. Exchange formerly covered a limited economic area. Cheap transportation (railways) enlarges the area; railways.—§ 6. Wider markets bring more minute division of labor. Illustration from butcher's trade.—§ 7. The geographical division of labor, illustrated by the United States and Great Britain.—§ 8. Two sorts of gain from geographical division of labor.

4. LARGE-SCALE PRODUCTION 47

§ 1. Growth of large-scale production closely connected with the Industrial Revolution. Advantages of large-scale production: use of machinery, saving in general expenses, buying and selling, utilization of by-products, experimenting.—§ 2. Limitations, chiefly

CHAPTER PAGE

from difficulties of superintendence. The case of agriculture. Other industries. The development in retail trade.—§ 3. Scarcity of able managers as a cause of limitation. This human factor usually ignored by the socialist.—§ 4. Large-scale management. Combination, horizontal and vertical. The Steel Corporation as an example. Other examples. The tendency to vertical combination less strong than that to horizontal.—§ 5. Competition often wasteful; tho the waste is less than it seems.

5. CAPITAL 62

§1. Production is spread over time. This fact disguised by the division of labor. Increasing use of plant and machinery in modern times.—§ 2. Producer's wealth and consumer's wealth; capital.—§ 3. Capital rests on a surplus.—§ 4. In what sense capital rests on saving. Hoarding contrasted with saving for investment.—§ 5. Investment means advances to laborers. Inequality of possessions in relation to advances. Middlemen for investment and advances.—§ 6. The maintenance of capital, as well as its creation, involves saving.—§ 7. Some difficulties of terminology.

6. CORPORATE ORGANIZATION 76

§ 1. Partnerships and corporations. Limited liability. Corporations from the legal point of view and from the economic.—§ 2. Advantages from corporate organization. Large-scale operations facilitated; new and venturesome investments promoted; stimulus to savings and investment.—§ 3. Ease of transfer serves to divide risks and so promote investments, and to bring control into capable hands. But it leads to great evils.

7. SOME FACTORS IN PRODUCTIVENESS 84

§ 1. The effect of high wages (abundant food) on the productivity of labor. High wages in the main a result, not a cause, of efficiency.—§ 2. Effects of skill and intelligence on productivity. General education. Technical education, in its effect for the individual and for the community.—§ 3. Leadership. The business man; the man of science. Freedom and mobility as promoting leadership. The motives to leadership.—§ 4. The immaterial equipment of a community; how affected by training and by inheritance.

REFERENCES ON BOOK I 96

BOOK II

VALUE AND EXCHANGE

8. INTRODUCTORY: EXCHANGE, VALUE, PRICE 99

§ 1. Exchange the consequence of the division of labor.—§ 2. Money as the medium of exchange.—§ 3. Value and utility. The

CHAPTER PAGE

notion of value in exchange.—§ 4. A general rise in values; a general rise in prices. Stability in general prices provisionally assumed.

9. VALUE AND UTILITY 106

§ 1. Utility a necessary condition of value; but value not proportional to utility.—§ 2. Increase of supply brings lowering of value; because of differences of means, and, fundamentally, because of the tendency to diminishing utility. Effects of varying the commodities supplied. Possible exceptions to the general principle. —§ 3. Total utility and marginal utility.—§ 4. Value depends on marginal utility. Qualifications and explanations. Marginal vendibility. The marginal utility of money.—§ 5. Consumer's surplus. Sundry limitations on its significance and on the possibility of measuring it.—§ 6. How state and measure the income of a community?—§ 7. The law of diminishing utility points to the conclusion that inequality lessens maximum well-being.—§ 8. Concluding remarks on welfare and total utility.

10. MARKET VALUE. DEMAND AND SUPPLY 124

§ 1. The conditions of demand and the demand curve.—§ 2. Demand possibly discontinuous, usually continuous. Elastic and inelastic demand.—§ 3. How value is determined by marginal vendibility, for a fixed supply. The equation of demand and supply. —§ 4. A varying supply: the equilibrium of demand and supply.— § 5. How far the supposition of a fixed supply, how far that of a varying supply, conforms to the facts. The circumstances that act on daily and on seasonal prices.—§ 6. Qualifications as to the market value of capital goods.—§ 7. Retail prices seem to follow wholesale prices, but in the end govern wholesale prices. The advantage of fixed retail prices.—§ 8. Current market prices are what people commonly mean when they speak of "fair" prices.—§ 9. Sporadic cases where value is affected by utility to sellers.

11. SPECULATION 145

§ 1. The fundamental effect of speculation by "legitimate" dealers is to mitigate fluctuations.—§ 2. Dealing in futures lessens price fluctuations of goods.—§ 3. Exchanges; standardizing.—§ 4. The evils of speculation, especially by "outsiders": gambling; unproductive labor.—§ 5. Stock exchange speculation.

12. VALUE UNDER CONSTANT COST AND FREE COMPETITION 157

§ 1. The simplest case first assumed: a supply absolutely flexible, free competition, constant cost. Value then determined by cost.— § 2. Illustration by diagram.—§ 3. The proposition points to a tendency or approximation only; to what happens in a "static," not in a "dynamic," state.—§ 4. Some explanations and qualifications. Flexibility in supply never perfect, often much impeded.

CHAPTER PAGE

Changes in demand from fashion. Good will. A small surplus above cost price may mean large profits.—§ 5. The extent and the importance of the divergence from competitive conditions.

13. VALUE AND VARYING COSTS. DIMINISHING RETURNS . . 166

§ 1. The equilibrium of value where marginal vendibility and marginal cost balance. The simile of the scissors.—§ 2. Permanent variations in cost affect long-run value differently from temporary variations.—§ 3. Diminishing returns.—§ 4. Permanent variations, or diminishing returns, appear most in the extractive industries.—§ 5. Something analogous as regards business management.

14. VALUE AND INCREASING RETURNS 175

§ 1. External economics.—§ 2. Internal economies. Large scale establishments.—§ 3. Effects of the two on industry at large. on the development of monopoly.—§ 4. What is meant by a "law" of increasing returns. How far there is a law as regards external economies; as regards internal economies.—§ 5, 6. A theoretical treatment of increasing returns. Successive or historical supply curves. The interaction, over long periods, of increasing and diminishing returns.

15. JOINT SUPPLY AND JOINT DEMAND 190

§1. Joint supply, or joint cost: effect of increase or decrease in demand. Influence of separable items of expense. "By-products." Complex case where both monopoly and joint cost exist. Influence of large plant.—§ 2. Joint demand. The constituent most limited in supply feels most the effect of an increase of demand. Labor in building trades as an illustration. Joint demand usually causes peculiarities less enduring than those arising from joint supply.

16. THE INDIVIDUAL FIRM 196

§ 1. The individual firm in the recent development of theoretical analysis. The firm's management of production here considered, not trading operations.—§ 2. The generalized principle of diminishing productivity, or law of varying proportions. Illustration from land, and seed applied to land.—§ 3. Other factors than seed, other industries than agriculture.—§ 4. The short-run cost curve of the individual firm.—§ 5. The short-run adjustment of output to maximization of profit. The long-run adjustment.—§ 6. The position of the individual firm different according as it makes a standardized or a differentiated product. In the latter case, the demand curve likely to be more steeply inclined. Marginal cost and marginal revenue.—§ 7. Conclusion. Bearing and importance of this type of analysis.

17. MONOPOLY VALUE 214

§ 1. Monopoly affects price thru limitation of supply. This proposition qualified as to transactions between middlemen, especially

CHAPTER PAGE

as to producer's capital.—§ 2. How price is fixed if a monopolist has a fortuitous supply; how, if he produces his supply at constant cost. Monopoly profit. Destruction of part of the supply possible, but not probable. Diamond mining as illustrating monopoly price. —§ 3. Monopoly price under increasing returns. Copyrighted books as illustrations. Monopoly price under diminishing returns.—§ 4. Possibility of varying prices under monopoly. Copyrighted books; telephone rates. Converse case of uniform prices under monopoly. —§ 5. "Dumping" explained by monopoly.—§ 6. Unqualified monopoly rare; various limitations and qualifications.—§ 7. Duopoly; complications of the theoretic problem; the complications even greater in the world of affairs. In what way and to what extent competition remains an active force.

18. SEMI-MONOPOLY. MONOPOLISTIC COMPETITION . . . 232

§ 1. Brands, trade-marks, advertising. Their good side tolerably clear as regards producers' goods.—§ 2. Less so as regards consumers' goods. Costs made higher.—§ 3. Effects of the advertising of branded goods. Difficult problems in the strict field of economics. —§ 4. Even more perplexing problems as regards welfare and happiness. Questions as to the scope of economics.

REFERENCES ON BOOK II 239

BOOK III

MONEY AND BANKING

19. INTRODUCTORY 243

§ 1. This Book takes up chiefly questions arising about the general level of prices. The value of money. What "money" means: coin, paper money, deposits.—§ 2. Coins and coinage. Gold and silver and their long use. Coins, formerly dominant, now in wide use for small payments only.—§ 3. Paper money and its use in modern times. Convertible and inconvertible paper.—§ 4. Bank deposits and the check system. The legal position of banks and of their depositors. Growth of deposits in English-speaking countries; why classed as "money."—§ 5. The "quantity theory" of money. The demand for money has an elasticity of unity.—§ 6. Explanations and qualifications.—§ 7. The way in which money and goods meet each other in the market. Velocity of circulation. Custom the chief determinant in the circulation of money.—§ 8. It is not important whether the price level be high or low, but highly important whether it changes.—§ 9. How reckon the total quantity of money at a given time and place?

20. CHANGES IN PRICES 259

§ 1. Changes in the price level are measured by index numbers. The arithmetical mean. Illustrations from U. S. prices, 1913–18.

CHAPTER PAGE

—§ 2. Other means: the geometrical, the median. Weighted index numbers.—§ 3. Prices and money, incomes; debtors and creditors. —§ 4. Complications. Changes in particular prices ordinarily are of chief direct concern. As regards the price level, abrupt changes alone have far-reaching consequences.—§ 5. Debts having a long time to run. The problem where the general trend of prices is different from that of money incomes.—§ 6. Prices and costs. Effects of changes in their relation resulting from the nature of capitalist industry. Business concerns and business men as debtors and creditors.—§ 7. Prices and costs further considered. Their relation is unevenly influenced by industrial oscillations.—§ 8. Wages as costs and their relations to prices. Similar problems with monopolized raw materials.—§ 9. Concluding remarks. The commodity dollar.

21. COIN AND SPECIE 281

§ 1. Gold and silver, long the dominant forms of money, gradually became of less importance in the 19th and 20th centuries. The qualities that led to their use as money.—§ 2. Their position strengthened by factors other than the physical qualities. Convention and habit.—§ 3. Coinage a public function. Under the gold standard, gold and bullion interchangeable.—§ 4. Bimetallism, and the difficulties to which it led. Its history. Silver suddenly discarded toward the close of the 19th century. The attempt to revive it in the U. S. from 1878 to 1934 an anomaly.—§ 5. Some problems of coinage. Subsidiary coin.

22. BANKING OPERATIONS. BANK NOTES 296

§ 1. Function of banks. Savings banks typical as channels of investment. Other similar banking agencies.—§ 2. Commercial banks; their customers and depositors. Their demand liabilities (bank notes and deposits) constitute part of the effective money supply. —§ 3. Notes of private banks, long important, have virtually ceased.—§ 4. The ready circulation of bank notes and their reckless issue during the first part of the 19th century gradually led to regulation and control.—§ 5. Complete security for holders of bank notes finally achieved in the United States. Should depositors have the same security?

23. DEPOSIT BANKING AND DEPOSIT MONEY 306

§ 1. Deposits may be lodged, the original form; or they may be created. Continuing deposits and their circulation.—§ 2. How long a created deposit remains effective as money. The relation between created and lodged deposits.—§ 3. Is the creation of deposits semi-fraudulent?—§ 4. Overdrafts in England.—§ 5. Clearing houses and the inter-relations of cash and deposits.—§ 6. Bank notes and government paper of the smaller denominations now play in English-speaking countries a part similar to that of subsidiary coin.—

CHAPTER PAGE

§ 7. Commercial banks and permanent investment. Command of capital precedes savings. Investment is promoted and hurried, with consequences that reach far, and not seldom are injurious.—§ 8. Changes in the character of banking and in the characteristics of bankers.

24. DEPOSITS FURTHER CONSIDERED. RESERVES, RUNS. REGULATION 325

§ 1. Cash in the vaults of a bank tends to be reduced to the minimum. But its resources must be largely "liquid"; commercial paper, saleable securities, demand loans.—§ 2. Relation between the rates of discount on bank loans and the reserves held by the banks. Demand loans and speculative transactions.—§ 3. Runs and panics.—§ 4. Far-reaching effects of panics.—§ 5. Repeated breakdowns in the American banking system. Peculiarities of monetary and banking conditions.—§ 6. Further remarks on the peculiarities of deposit money.—§ 7. Regulation for security of deposits in the United States. The great step, insurance, taken in 1935.

25. DEPOSIT BANKING, *Concluded* 343

§ 1. The way in which the general welfare is promoted by commercial banking operations: continuity of industry, selection of the more capable among business men.—§ 2. Investment operations of commercial banks. Deposits lodged by persons not active in business.—§ 3. Creation of deposits and the march of progress. Criticisms of the deposit structure and proposals for its radical change.

26. CENTRAL BANKS; FRANCE, ENGLAND 348

§ 1. Three great systems to be examined, those of France, England, and the United States. In all three countries, the Central banks, while private corporations in form, are in effect agencies of the government.—§ 2. The Bank of France has a complete monopoly of note issue. Importance of this, in view of the comparatively slight development of deposit banking.—§ 3. Wide permeation of the Bank's operations. Rediscounting and the effects. The Bank essentially a banker's bank.—§ 4. Close connections of the Bank with the public finances. In times of stress, an agent for putting out inconvertible paper money.—§ 5. The Bank of England. Its organization. Absence of branches. Complete separation of notes and deposits. Issue and Banking Department. The Issue Department and the fiduciary issue. Bank notes and currency notes.—§ 6. The Banking Department. Relation to the other banks. The Bank's reserve. Bill brokers.—§ 7. Working of the system in times of crisis, use of the discount rate, open market operations.—§ 8. The Bank and the government. Enlarged note issue and inflation in 1914–18. New functions after the war, and especially since 1931. Growing concern for industrial as well as monetary stability.

CHAPTER PAGE

27. CENTRAL BANKS, THE FEDERAL RESERVE SYSTEM . . . 363

§ 1. The old National Bank System. The Federal Reserve System; the Board of Governors and its powers.—§ 2. Note issue by the Reserve Banks only; Deposits in the Reserve Banks. Reserve requirements for the Reserve Banks.—§ 3. Working of the System in times of panic.—§ 4. Reserve requirements for the Member Banks. Flexibility of the system in its more recent phases.—§ 5. Rediscounting. —§ 6. Other instruments of control in the Board's hands: changes in the rate of discount; open market operations, power to alter reserve requirements of Member Banks.—§ 7. Character of the system as a whole, and its future.—§ 8. Central Banks in general. Sole issuers of paper money. New and growing functions and responsibilities.

28. GOVERNMENT PAPER MONEY 378

§ 1. Paper money has come to be that put out directly or indirectly by governments.—§ 2. Convertible notes.—§ 3. Inconvertible paper. How its value rests on its quantity.—§ 4. Extent to which paper has displaced coin. Its convenience. Custom as affecting the proportion of paper to the total of active money.— § 5. Inconvertible paper may operate in two ways: as hand-to-hand money, and as the basis of the entire monetary structure.—§ 6. How far government can maintain the acceptability of inconvertible notes.—§ 7. The effects of sudden large issues; extreme depreciation and collapse.—§ 8. Experiences during the World War of 1914–18.—§ 9. Unexpected consequences of these experiences in countries that adhered to the gold standard.—§ 10. The specie premium.—§ 11. Problems arising when there is return to the specie standard.—§ 12. Great issues of inconvertible paper are almost invariable consequences of prolonged war.

29. THE THEORY OF PRICES IN REVIEW 398

§ 1. The Quantity theory once more. The term "money." How far gold remains "money."—§ 2. Barter transactions, and others settled without use of cash. Credit where there is mere postponement of payment. "Bank credit," as this term is often used, raises different questions.—§ 3. How far the two factors—volume of money and volume of goods—are independent. Effects of changes in price on volume of production.—§ 4. Effects of changes in production on prices. Deposits when created in conjunction with changes in transactions. Acceleration of prices in the upswing period of a business cycle.—§ 5. The bank rate of interest and the price level.—§ 6. The case of government paper.—§ 7. What remains of the Quantity theory?

30. THE GOLD STANDARD 414

§ 1. The essentials of the gold standard: the monetary unit convertible into a given weight of gold. Its older form: free coinage

CHAPTER PAGE

and free circulation.—§ 2. The later form: the gold bullion standard. Contrast to free circulation. Relation to paper money. A variant: the gold exchange standard.—§ 3. Two gains ascribed to the gold standard; stabilization of the foreign exchanges, and stabilization of the price level. The first clearly achieved; its importance sometimes exaggerated.—§ 4. Stabilization of prices more important but never so fully attained. How affected by the form of the gold system, by the sensitiveness of the banking system and by the character of international payments.—§ 5. Difficulties after the War of 1914–18. England. The United States. Other countries.—§ 6. The only alternative to the gold standard system is one of inconvertible paper. Both call for management but the paper for more of it. Importance of psychology, of rooted traditions; of political factors and international relations. Some conclusions.

REFERENCES ON BOOK III 433

BOOK IV

INTERNATIONAL TRADE

31. THE FOREIGN EXCHANGES 437

§ 1. The "foreign exchanges," based on the varying coinage systems of different countries. In this chapter considered on the supposition that the gold standard prevails thruout. How bills of exchange settle payments without the movement of specie.—§ 2. The par of exchange, and premium and discount of exchange; illustrated by sterling exchange in New York.—§ 3. Bankers as middlemen in the foreign exchanges. Fluctuations in rates, due to the higgling of the market.—§ 4. Dealings between a series of countries, illustrated by transactions between the United States, England, and Brazil. The wide use of sterling bills for trade between all parts of the world.—§ 5. How prices of bills are influenced: in the long run, by the flow of specie; for shorter periods, by the rates of discount. Various complicating factors.

32. THE BALANCE OF INTERNATIONAL PAYMENTS 452

§1. Other items than merchandise exports and imports. Lending and borrowing and their effects on exports and imports. International dealings in securities.—§ 2. Expenses of travelers and nonresidents. Remittances from the United States by immigrants. Freight charges.—§ 3. Position of a country that mines specie.—§ 4. Illustration from the international trade of the United States, 1790–1908.—§ 5. The notion of a favorable and unfavorable balance of trade. Usual attitude of the business community. In the main, an excess of imports or of exports is no indication of loss or gain; least of all, in the trade between one country and any other

CHAPTER PAGE

country.—§ 6. Unusual character of the international trade of the United States during the World War 1914–19; and again in the post-war period.

33. DISLOCATED EXCHANGES 469

§ 1. Exchanges become dislocated when the currencies of the trading countries are not on the same basis, e.g., not on the gold basis.—§ 2. Effects of dislocation in the simplest case: when merchandise transactions are the only ones between countries, and prices in each country are stable in terms of its own currency.—§ 3. The case usually is complicated. Illustration: changes in tourist expenditures; effects on imports and exports of goods; "bounties" on imports or exports.—§ 4. Prices of domestic and of international goods do not move together.—§ 5. Dislocations arising from changes in the monetary standard itself. The United States in 1933–36.—§ 6. Effects of complete abandonment of the gold standard, "Exchange dumping."—§ 7. The effects of dislocated exchanges are of the short-period kind. In the long run the character of a country's international trade is not altered by them.

34. THE THEORY OF INTERNATIONAL TRADE. WHY PARTICULAR GOODS ARE EXPORTED OR IMPORTED. THE PRINCIPLE OF COMPARATIVE ADVANTAGE 481

§ 1. Some familiar facts: money incomes and prices differ in different countries; but prices of goods entering into international trade tend to be the same. Money wages not necessarily low in exporting countries.—§ 2. A country exports those things in which its labor is relatively effective—in which it has a comparative advantage. Illustrations from countries of high wages and of low wages.—§ 3. Differences in advantage may rest on climate or other physical cause, or on acquired aptitudes.—§ 4. Specially low wages of a particular class of laborers operate in the same way as does a comparative advantage in real cost. General low wages do not affect international trade or enable universal underselling.—§ 5. A country may import things for which its labor is productive, if its labor is even more productive for other things. But international trade rests largely on absolute differences.—§ 6. The gain from differences in comparative cost is dependent on immobility of labor between countries.—§ 7. A country may import part of the supply of a given commodity, produce a part at home. Difference between extractive and manufacturing industries in this regard.

35. THE THEORY OF INTERNATIONAL TRADE, *Continued.* WHEREIN THE GAIN CONSISTS 495

§ 1. Difference between trade within a country and international trade. Varying rates of wages in different countries show varying gain from the exchanges between them.—§ 2. An illustrative case;

CHAPTER PAGE

England and Italy. Demand and utility determine relative wages and prices. Slow and obscured operation of this cause, thru the influence of the money supply on incomes and prices.—§ 3. Effects of changes in international demand; of new articles of export; of payments other than for merchandise.—§ 4. Difficulty of following these causes in detail, illustrated by the case of the United States since 1873.—§ 5. Money incomes, not prices, important in determining the gain from international trade.—§ 6. Two causes act on the gain: the play of international demand and the effectiveness of labor in producing exported goods. The last cause settles the general rate of money wages.—§ 7. High money wages and other incomes do not necessarily bring high domestic prices. Illustration from the United States.

36. PROTECTION AND FREE TRADE. THE CASE FOR FREE TRADE 508

§ 1. The main argument for free trade is simple. Persistence of Mercantilist notions.—§ 2. Some popular arguments for protection: creating a home market; the truck farm case; creating employment.—§ 3. Protective duties as a means of combating industrial depression.—§ 4. The effect of protection on wages. Real wages in general lowered, tho some particular wages possibly kept up.—§ 5. The principle of equalizing cost of production.—§ 6. Effects of duties on prices and on consumers. A national loss only if domestic products are substituted for those imported. Monopoly may bring special gain to domestic capitalists, but brings no national loss. Labor monopoly may bring special gains to particular laborers.

37. PROTECTION AND FREE TRADE, *Continued.* SOME ARGUMENTS FOR PROTECTION 525

§ 1. Protective duties, by their effects on general money incomes, may bring more advantageous terms of international exchange.—§ 2. Protection to young industries. Applicable in the main to manufactures only. Difficulty of gauging its success in specific cases.—§ 3. Political considerations, illustrated by the case of shipping subsidies.—§ 4. Subsidies, much resorted to in the 17th and 18th centuries; less so in the 19th; more so again in the 20th.—§ 5. Social considerations may tell against manufactures, but not necessarily so. The controversy in Germany; Agrarstaat *vs.* Industriestaat. The argument as to the failure of food supplies.—§ 6. Peculiar dependence of England on international trade and on exports. Possibility of strengthening her position as exporter by agreements with colonies and by threats of retaliation.—§ 7. Growth of protection during the 20th century. Still greater growth after the War of 1914–19. Autarkie in Germany; in the British Commonwealth.—§ 8. Economic effects of protection in the United

PAGE

States; impossible to measure accurately, but certainly exaggerated in popular discussion.—§ 9. Conditions under which manufacturers would maintain themselves without protection. Effect of machinery in connection with comparative costs. Concluding remarks on the working of protection in the United States.

REFERENCES ON BOOK IV 546

BOOK I

THE ORGANIZATION OF PRODUCTION

CHAPTER 1

WEALTH AND LABOR

§ 1. The subject matter of economics.—§ 2. Wealth; free goods; economic goods. Wealth and welfare.—§ 3. Goods may become economic thru mere scarcity; but commonly do so because of the exertion of labor.—§ 4. Activity may be irksome or pleasurable. Labor usually is continuous, little varied, irksome.—§ 5. Some sorts of labor always pleasurable. Economic welfare: happiness.—§ 6. The irksomeness of most labor to be lessened by better general opinion, and by greater leisure thru shorter hours.

§ 1. To define with accuracy the scope and contents of economics is not of importance in the earlier stages of its study. The precise demarcation of its subject matter, and its relation to other branches of knowledge, can be understood only when something is known of its main conclusions. It suffices at the outset to indicate by an example what is the nature of the problems dealt with. A good example is found in the economic position of one of the most familiar articles of use—water.

In a thinly settled community, where springs and streams are abundant, water is free to all. No question can arise as to its ownership or as to the mode in which the community should deal with it. Every one is fortunate in having an unlimited supply. No one can gain advantage by taking possession of part of it, or devoting labor to procuring it.

Water under such conditions is said to be a "free" good, not an "economic" good. It is not an economic good, in the sense that no economic problems arise regarding it. Every one has all he wants, and thereby is prospered; what more is there to say?

A stage may come very early when some labor will be given to making the water conveniently available, and when it will be no longer strictly a free good; and when yet no economic questions of any complexity arise. The individual may dig a well, or pipe the water from a spring or stream to his dwelling. The very first economic problem, that which may even be considered the funda-

mental problem, then emerges: How much effort is it worth while to give to the supply of this convenience? But the problem remains a very simple one so long as the individual exerts himself to satisfy his own wants only. There is no dealing with others, no sale, no question of price. If men were to work solely for the satisfaction of their own wants the difficult economic questions would not arise.

A more complex stage is reached when water is brought in by some individuals and sold to others. In oriental towns the water carrier, with his runlet or skin, is still a familiar figure. In our own cities private individuals sometimes sell carboys of spring water or distilled water. Here questions of sale and price arise. What settles the terms on which water is sold? What settles the earnings of those who supply it? Are they in a position of advantage or not? Here are matters less simple.

Still another stage (not necessarily a later stage) is reached when common action is taken to procure the water. Here the problem may remain comparatively simple, or it may become one of the troublesome problems of modern communities. The traveler in Italy sees the village fountain, supplied by its aqueduct; even in larger towns, thru some parts of Europe, the public fountain has remained until very recently the chief source of supply. The water is no longer strictly a "free" good, since effort and expense were required to bring it where wanted. But the effort was made long ago, does not need to be renewed (there are no expenses of upkeep), and there is so much water that it can be used without restriction or regulation. In the modern city, however, the case has become different. There are great reservoirs, elaborate pumping stations, mains and pipes. Water is supplied abundantly and conveniently to every household. There is not only a vast initial outlay for the plant, but a continuing cost of upkeep. The questions arise, Who shall make the outlay and manage the supply? Shall there be public or private ownership? And, whether under public or private ownership, what are to be the conditions of sale? The water, if under public management, may still be supplied gratuitously to all, as it is at the village fountain; or payment may be required of the users. The questions of price arise, of sound

public policy, of possible gains to the public treasury, of conflict between financial and hygienic considerations. The really complex problems of economics arise full-fledged.

§ 2. To designate these different sorts of conditions, some quasi-technical terms are often used: "free goods," "economic goods," "public goods," "wealth."

What are free goods and what are economic goods has just been indicated. Fresh air, climate, sunshine, are the obvious cases of free goods; so is water under the simplest conditions, or standing timber in a thinly settled and well-wooded country.

Scarcity is the earmark of an economic good—scarcity, that is, relatively to the demand. Water becomes an economic good when effort is needed to obtain it in the quantity desired at the place of use. Conceivably fresh air will become in the future, for considerable numbers of mankind, an economic good. It is so already when many persons are gathered in a large room or hall. Fans, conduits, engines, are installed; it becomes a question how the needful efforts shall be best directed, who shall bear the expense. With the concentration of population in great cities, and the multiplication of agencies that pollute the air in them, it is possible that elaborate means will have to be taken for keeping it healthful. Then the same complex problems will present themselves as in the case of water; all resting on the relative scarcity of the thing in question.

"Public goods" are economic goods supplied gratuitously to individuals, yet involving effort and consequent expense to some one. Tho free to the users, they are not free goods. Such is water at the public fountain; such are parks, museums, free concerts, public education, bridges, and highways. What goods shall be public, and by whom the expense of providing them shall be met—whether by levy on all persons, or on some only—these are problems as to public functions, and as to taxation for defraying their expense, among the most difficult and far-reaching that the economist has to deal with.

It was common in the older books on our subject to define political economy, (a phrase replaced in modern times by the simpler "economics") as the "science of wealth." In this usage, wealth meant all the economic goods, including the public goods.

Either term—wealth or economic goods—serves to describe the subject matter with which economics has to deal; those things which men want, which are not free, and which present the problems of effort, of satisfaction thru effort, of the organization of industry.

Evidently a community is the better off, the more free goods it has and the less the range of things that come within the category of "wealth." Where unlimited pure water and fresh air are at every one's disposal, the conditions of life are eased by so much. A mild and equable climate relieves the people of some favored spots from much labor that must be given elsewhere to protection from heat or cold. It may be said, with an appearance of paradox, that the more things in the nature of wealth a community has, the less prosperous it is. The paradox is easily solved. The wealth of a community is not the sum total of things on which its welfare depends—these include its free goods as well as its economic goods. The more things are free, the easier are the conditions of living. The more things are economic, the wider is the range of commodities concerning which the economic problems arise, and the wider is the scope of the science of "wealth."

The abundance of free goods, tho in itself advantageous to a community, does not always coexist with the highest degree of prosperity. In tropical and semitropical countries the conditions of living are on the whole easier than in temperate countries. Some sorts of food are free or nearly free, and protection does not need to be provided against the cold of winter. But the climate saps energy and checks the development of physical vigor and of intellectual capacity. Hence the peoples of temperate regions, from the very obstacles they have to overcome, gain resources within themselves which lead eventually to greater prosperity. So it is with individuals. He who has always had abundant means at his command often lacks endurance and spirit, and in the end is surpassed in happiness as well as in riches by him who had to face harder conditions at the start.

§ 3. In the preceding paragraph wealth has been spoken of as the result of effort. But there are cases where a commodity is wealth—is an economic good—even tho it be obtained without

effort. A free gift of nature may be wealth, if it is limited in quantity.

Meteoric stones, usually disintegrated by heat before touching the earth's surface, in some instances reach the ground. Being scarce, and in our days esteemed for scientific research or even for the satisfaction of mere curiosity, they command a price. Tho the free gift of nature, they are not free goods in the economic sense. On some parts of the seashore the waves dislodge from near-lying rocks quantities of kelp, which is useful as a fertilizer. Like multitudes of other articles, its use is indirect; it does not satisfy wants directly, but is an aid in the operations for satisfying them. Obviously, it may none the less be wealth. If kelp were steadily swept to the shore in such quantities that every one could get all he wished, it would be a free good in the strict economic sense. But if it is deposited in limited quantities on favored spots, and if many farmers are desirous of using it, it will command a price as it lies on the beach, before the hand of man has touched it. And the same quantity which at one time was so abundant as to command no price may be brought by the growth of population within the circle of things bought and sold, and so become one of the goods with which economic science deals.

The same narrowing of the circle of free goods and the same widening of that of economic goods or wealth appear if there be not a natural but an artificial scarcity of goods. A supply of water or timber, unlimited in quantity for the needs of a given community, may come by force or by long-settled law under the control of some individual or individuals. By limiting the amount which others shall have, the owners may make such things a source of income for themselves and cause them to enter the list of economic goods. Monopoly by itself raises some of the questions with which economic science has to deal.

This simplest sort of scarcity may seem to be exceptional; and as to the things which we usually think of as goods or commodities it is so. The instances just adduced are exceptional. In the vast majority of cases commodities become economic after some labor has been applied to fashioning them. Tho scarcity (that is, relative scarcity) still underlies the notion of wealth or economic

goods, it is scarcity in the sense that the materials supplied by nature need to be adapted to man's use by his labor. Labor, or effort of some sort, is usually the cause or condition underlying economic phenomena.

There is one large class of things, however, for which this statement does not hold: limited natural agents, of which land is the most conspicuous. These are not commonly called goods or wares; but they are economic goods in the strict sense, being limited in quantity and of high service in satisfying wants. Agricultural land, power and deep-water sites, forests, mineral lands,—all are often economic goods by virtue of mere natural limitation of quantity. They present, as will appear in due course, some of the most intricate social and economic problems.

§ 4. What constitutes labor may seem a simple matter. Most people would say that they are more than sufficiently familiar with it. Yet some questions arise concerning it that go to the heart of economics, and the last word on them cannot be said until the very close of the exposition of the whole subject.

Some activities are agreeable, some are irksome. Some are undertaken for the pleasure of doing, some for a reward. Not infrequently the two satisfactions are gained simultaneously from the selfsame activity; it is both a source of pleasure in itself, and it brings a reward.

So far as the nature of the muscular or nervous effort is concerned, no distinction can be drawn between the agreeable and the irksome activities, or between those which are undertaken for pleasure and those which are undertaken for pay. Such severe physical labor, combined with hardship and exposure, as mountain climbing is done for pleasure by tourists and for pay by guides. The pursuit of athletic sports is the most familiar of recreations and is also a familiar profession. A multitude of occupations ordinarily pursued for gain—woodworking, gardening, painting, acting—are also pursued by many persons for the satisfaction which the doing affords.

None the less it is true that the greatest part of the activity which men carry on in getting a living does not give pleasure. The chief reason seems to be that activity, in order to be effective

toward getting a living, must be steady, unvaried, and long-continued; and it must be, in an important sense, not free. The characteristic of most activities that are sources of pleasure in themselves is the element of freshness or novelty, and the absence of any sort of compulsion. The guide who climbs mountains year after year, and knows the tracks by heart, soon finds the task a weary one; and this the more because in order to earn his living he must follow his tracks regularly, regardless of health or spirits at the moment. It is the zest of novelty and the sense of freedom and choice that bring pleasure in the summer's strenuous vacation. Inactivity and idleness soon become irksome; but, with few exceptions, steady application to the same task also soon becomes irksome.

In savage and barbarian communities, the men usually confine themselves to the chase and to war. The monotonous work of cultivating fields and of preparing food is left to the women. Tho hunting and fishing often entail the severest hardship, they do not commonly endure long and they are almost surely varied by changes and respites. The variety and the sudden changes give play for emulation and for satisfying the love of distinction, (and that for slaughter also) impulses which have a powerful effect in many fields of economic activity. An alternation of periods of complete idleness and of feverish activity is characteristic of those early stages of society in which men give themselves to the unchecked satisfaction of their propensities.

The sort of labor that occupies the mass of mankind in civilized societies, and that which brings the largest product, is mainly of the continuous, monotonous, and usually irksome kind. This is more especially the case where the division of labor has been much elaborated. The wide extension of the division of labor, which has been a main cause in modern times of the greater abundance of material goods and of the extraordinary advance in material prosperity, has probably also been a cause of greater weariness and unattractiveness for most labor. Even in the simpler and older form of the division of labor, where one man was carpenter, another smith, another cobbler, there was of necessity a steady repetition of operations and no little monotony of work.

In the remarkable splitting up of occupations which has resulted from the elaboration of machinery in modern times, it is rare that a workman does all the work of his trade or even knows how to do it. He is no longer a cobbler making a whole shoe but a factory hand attending hour after hour and week after week to the same specialized piece of machine work. Moreover, in a dense population and with strictly enforced ownership of property and of land, he is under compulsion to do continuous work of some such sort; he must do it in order to keep body and soul together. He lacks variety and he lacks freedom. He may find pleasure in exerting himself strenuously at sports but the labor of getting his living yields in itself little satisfaction.

§ 5. Some sorts of labor, tho pursued systematically and continuously, seem never to become wearisome. This is the case with much intellectual labor, especially that of persons who are engaged in the pursuit of knowledge and in the satisfaction of man's insatiable curiosity about the things that surround him. Persons of artistic temperament—painters, musicians, poets—have often so strong an inborn bent toward one kind of activity that nothing can hold them from it and nothing ever pall the pleasure of the exertion. And any occupation which satisfies the instinct of emulation has unceasing charm. He who can achieve things which few can achieve and which many would like to achieve, rarely tires of his work. The actor, even tho his occupation involves the monotonous and long-continued repetition of the most trifling details, never fails to get a thrill of pleasure from the breathless silence or stirring applause of his audience. Were he compelled to go thru his part as often and as rigorously under the cold supervision of an indifferent supervisor, and under that only, how flat and stale it would become! For a similar reason, work of leadership and command almost always is continuously enjoyed. It satisfies the love of distinction and the desire for domination and it has a real or apparent element of freedom. Hence the work of the employer commonly affords more satisfaction than that of the employee, and often is continued, from mere love of the doing as well as from habit, long after the reward from the exertion has ceased to be in itself a source of much pleasure.

These exceptions should not blind us to the fact that by far the greater part of the world's work is not felt to be pleasurable. Some reformers have hoped to reach a social system under which all work would be in itself a source of joy. It is probable that such persons are made optimistic by the nature of their own doings. They are writers, schemers, reformers; they are usually of strongly altruistic character, and the performance of any duty or set task brings to them the approval of an exacting conscience; and they believe that all mankind can be brought to labor in their own spirit. The world would be a much happier place if their state of mind could be made universal. But the great mass of men are of a humdrum sort, not born with any marked bent or any loftiness of character. Moreover most of the world's work for the satisfaction of our primary wants must be monotonous, and often rough and coarse. There must be delving and ditching, sowing and reaping, hammering and sawing, and all the sustained physical exertion which, however lightened by tools and machinery, yet can never be other than labor in the ordinary sense of the term.

Much labor, then, there must be; not only hard but monotonous. It is not clear, however, that in these regards matters have become worse in modern times. Much of the severest labor—lifting, digging, pounding—has been taken over by machinery. Monotony may seem to have become greater under the influence of growing use of machinery and growing specialization of labor. But the effect of these changes on the happiness of living can be easily exaggerated. Ruskin has dwelt on the charm of the task of the medieval craftsman, who felt the joy of work that had beauty and character. Yet this joy was probably shared by few in medieval times or in any other. Then, as now, most work involved the repetition of the same operations, and was felt to be exacting and unwelcome. It is not easy to picture the conditions of life in earlier societies, organized in a very different way from our own; but it is more than probable that the mass of mankind found their tasks no pleasanter or lighter then than now.

All persons become aware, at one time or another, of a difference between the more immediate and the less immediate satisfactions from their daily doings. Tasks which at the moment may

be unwelcome are missed when omitted and are welcome when taken up again. Variety is pleasant, but pleasant only when it *is* variety. The life of the idle rich is often an unceasing round of endeavors to escape from boredom, alleviated (how much and how long?) if there is a feeling that most people envy the life and regard the idlers as distinguished personages. Individuals obviously differ; some are temperamentally Bohemians, others steady plodders. A life of simple living and serene contemplation is the ideal to some; one of unceasing and strenuous achievement to others; to still others one of alternation between intense activity and do-nothing vacancy. Most men are content when their lives ordinarily run in regular grooves, varied by occasional breaks, and become restless and discontented if they must make up their minds from day to day and from hour to hour what next to do.

Happiness, as these reflections indicate, is a different thing from economic welfare,—from the gratification or pleasure which comes for the time being from a given course of action. Happiness is a state of mind about which philosophers, psychologists, poets, preachers have much to say. Being quite subjective, a matter of individual feeling, it would seem impossible to say wherein it consists for mankind at large. Economic welfare is a narrower notion,—it refers to our feelings or satisfactions so far as dependent on the material things about us and chiefly on the results of the work which we do. These results, in the form of the goods and services we are willing to pay for, are of very great concern to all mankind. The task of economics is to examine the ways in which they become at our disposal and are apportioned among us: an important contribution to the understanding of well-being and happiness, even tho by no means the whole story.

§ 6. We may hope that as the material conditions of mankind improve, especially in the countries of advanced civilization, gains will be achieved as regards the irksomeness of ordinary labor. Some alleviation will come from a mere change in the state of opinion in the community. The sense of distinction affects the satisfaction from exertion. A task admired is an attractive task, and one despised is unattractive. The common attitude of the more favored classes has long been to contemn manual

labor and those who perform it. Such was the natural attitude in communities based on slavery or on its successor, feudalism, and such remains too often the attitude of that leisure class which in modern times apes many of the traits of feudalism. The growing democratization of society may be expected to change this and to raise the dignity and self-respect of labor of all kinds, manual or mental. Greater ease of movement between different classes and greater equalization of their conditions will add to the esteem in which all kinds of manual labor are held, and may remove some at least of the causes that now contribute to make it unwelcome.

The chief way, nevertheless, in which labor is likely to be made less irksome is not by a change in its character or its intrinsic attractiveness but by a diminution in its severity. It will probably be lightened by the increasing perfection of tools and the increasing use of machinery. More important is the prospect that the hours of labor are likely to be shortened, and the hours for recreation and variety correspondingly lengthened. The weariness of labor is by no means in proportion to the number of hours spent on it. For a healthy and well-nourished person, the first hours of exertion are not a source of fatigue. Some writers have indeed maintained that during these earlier hours—barring perhaps a brief initial period of stiffness—there is a sense of pleasure rather than of pain. This may be the case in intellectual activity and in some handicraft occupations, and the experience is a familiar one in holiday jaunts. But little direct consciousness of pleasure comes at any stage from the stated work of the great majority of men. The difference between the earlier parts of their day and the later is not so much that the former are pleasant and the latter unpleasant, as that fatigue does not begin until some hours have passed and then becomes increasingly severe with each of the later hours. When indeed the hours of labor are unduly prolonged, fatigue becomes so great and so deep-seated that the period of rest and sleep does not suffice to remove it. The next day begins again with fatigue, and worse succeeds worse. Such was the effect of the factory system in its early stages in England; such is still the situation in backward countries. Under these wretched conditions, the work of the day has covered eleven, twelve, even fourteen,

hours. In the United States, some of the steel-making industries, whose operations go on night and day, long had two shifts in each of which the men worked twelve hours. In such industries the substitution of three shifts and the reduction of hours in each from twelve to eight bring immense progress toward a life having possibilities of happiness.

The movement for shorter hours has been one of the most beneficent aspects of the betterment of material conditions in civilized countries during the last two or three generations. The day's labor was first cut down to eleven and ten, partly by the pressure of workmen's organizations and partly by legislation restricting the hours of women and children employed in factories. It is still in process of being reduced. The ideal of the trade unions long was to lower it to eight hours, a limit which has already been passed. We shall consider the significance of the shortening of the period of work, the causes of the gains so secured, and some fallacies which have attached themselves to the short-hour movement.[1] In itself that movement should have the hearty sympathy of every friend of humanity.

Notwithstanding all the alleviations of the irksomeness of labor —thru moderate hours and moderate tasks, free time for recreation, a general respect for labor of all kinds—the larger part of the world's work will always be felt to be irksome. A fortunate minority may work at tasks which are in themselves pleasurable and are not performed chiefly for the return which they bring. But most work is now undertaken for reward, would not be done without reward, and is strenuous and well directed in some relation to the reward. It is doubtless true that the mass of mankind, tho they find their labor irksome or repellent, are yet happier than they would be under complete idleness, or with only that fitful kind of exertion which attracts the savage. But labor is commonly felt to be a hardship, and the pay which it secures is the dominant motive for undertaking it. The problems that arise in economics are mainly concerned with the relation between exertion and the remuneration which induces the exertion.

[1] See Vol. II, Chapter 59.

CHAPTER 2

LABOR IN PRODUCTION

§ 1. Labor given to material objects deemed alone productive by earlier English economists. Objections to this view.—§ 2. Labor creates utilities only; all labor that issues in utility is productive. Is there nonmaterial wealth? —§ 3. Is there any unproductive labor? Labor given to things harmful.— § 4. Labor of judges and lawyers; of soldiers.—§ 5. Predatory labor. "Business." The law and unproductive labor.

§ 1. THE relation of labor to production may seem simple. Yet it has been the occasion of great differences of opinion among acute thinkers and it presents some nice questions.

We commonly speak of a tailor as making clothes, a carpenter as making a table, a cobbler as making boots. The familiar phrase, like most such, is elliptic and it leads easily to misunderstanding. The labor of the tailor gives only the finishing touch to the work previously done by a long series of persons—the shepherd who tended the flocks, the wool shearer, those who transported the wool by land and sea, the carder and spinner and weaver, not to mention those who made the tools and machinery of these workers. Similarly the carpenter is the last of a succession of persons who worked toward a common end—the lumberman in the woods, the sawyer in the mill, the trainman and the engineer on the railway, and so on. Many laborers, arranged in long series, combine in making even the simplest commodities.

But it is clearly all these laborers, taken together, who produce the commodities; and can it not be said these alone are the producers of wealth? Wealth has been described as consisting of those goods which are not free. The term refers primarily to things that are tangible and material. Many laborers produce no wealth in this sense. Such are domestic servants, policemen, actors, singers, teachers. Does not their work stand in a different relation to production from that of laborers who make material things and carry on production in the common meaning of the word?

This was the opinion of many of the earlier writers on eco-

nomics, especially the English writers from Adam Smith to John Stuart Mill. Their view was that only such laborers as turned out material things were productive; all others were unproductive. A liberal interpretation was indeed given to their definition of the productive laborers. Not only those who directly handled materials and fashioned them were included—the day laborer, the carpenter, and the smith, but those also by whom the operations were guided and promoted—the employer who directed the manual laborers, the foreman and the engineer, the teacher who trained the engineer. Even the teacher of the humblest workman may conceivably be regarded as contributing to the operations of material production in so far as the diffusion of even the rudiments of education raises intelligence and adds to efficiency. But with the widest latitude in interpretation, a great range of persons, doing all sorts of work and by it earning a living, remained outside the class of the so-called productive laborers. Domestic servants, lawyers and judges and policemen, all the army and navy, not to mention persons who provided mere amusement, were classed as unproductive. As Adam Smith remarked, "in the same class [of unproductive laborers] must be ranked, 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."

This distinction between productive and unproductive laborers was early attacked and long debated. It was pointed out that it seemed to affix some sort of stigma—an accusation of uselessness, of being in need of support from others—on whole classes of persons whose work was admitted to be honorable and often seemed to be indispensable. But this was after all not to the point; whether or no an "unproductive" occupation is to be regarded as honorable, the essential question is whether there are differences between this kind of work and the other which are important for the welfare of the community. It was much more to the point that the distinction led to difficulties and inconsistencies. The musician was regarded as an unproductive laborer; was the artisan who made his instrument—his violin—nevertheless productive? The labor of the violin maker issued in material wealth, or,

as Adam Smith said, in "a vendible commodity." Yet its only object was to make an instrument to be used by the musician; and was not the consistent view that of regarding the two sets of persons as combining for a common result, just as the sheep shearer, the weaver, and the tailor combine in making clothing? And if thus working together for the same end, was one to be set apart as productive, the other as unproductive? All members of the navy and army were classed as unproductive; yet those who built the ships, made the guns and the powder, were supposed to be productive. If one set were unproductive, why not the other?

§ 2. The solution of these difficulties is indicated by a conception which the British economists, tho they followed it in other directions, were curiously slow to use with reference to their discussion of productive labor. It points to satisfactions, or "utilities," as the aim and end of production. We shall see as we progress how in various directions economic science gains, and is often brought to unity and consistency, by the analysis of production as ending in utilities.

If it is a misleading use of language to speak of a carpenter as "making" a table, it is also misleading, tho in a different way, to speak even of a group of associated workers as "making" anything at all. The lumberman, the sawyer, the railway crew, the carpenter, between them are impotent to add to the amount of matter in the world. All that man can do is to change forms and combinations. And just this he does; he fashions and refashions material things. He puts them into forms in which they serve his wants. Such is obviously the nature of the carpenter's work, the tailor's, the cook's. It is not less true of those whom we describe as producing raw materials. The plants from which man secures the greatest part of his food and most of the material he uses get their constituent parts from the soil and the air. What man does is to arrange conditions favorable for their growth. The minerals which he uses are a fixed store in the earth's crust. When we say that coal is produced, we mean that it is brought to the surface and made available for our use.

The ways in which man brings about utilities or satisfactions

are many. Not only are plants grown, and coal, iron, copper brought up from the mines; not only are these raw materials shaped and adapted for their different uses—they are also transported to the places, often very distant, where they reach the hands of those whose wants they finally satisfy. They are bought by traders from one set of persons and sold again to another; and among the traders there is a division of labor, some buying at wholesale and selling again to the retailers, who in turn dispose of the commodities to their customers. The phrase "place utility" has been used to describe the contributions of those engaged in transportation and trade; and it serves to bring into relief the fact that such persons, tho they do not shape or fashion commodities, yet contribute to their utilization.

Now, since the essence of production is that it leads to satisfactions or utilities, it follows that any labor or effort that yields utilities is productive. The musician whose performance brings us pleasure does precisely the same sort of thing as the florist whose blossoms last a few hours. The domestic servant contributes to our ease just as does the artisan who supplies the furniture for our dwellings. No doubt there are gradations in the importance of the wants supplied by different workers. The essentials of life are most important; the conveniences and luxuries, the apparatus of display, come after them; and these gradations, as we shall see, have economic consequences. But they are not significant for the present purpose; they give no ground for distinguishing between those producers who embody utilities in material objects and those who do not. If we were called on to dispense with the services of some of the producers, we might put aside, as easily spared first, the buffoons and the opera dancers who figure as unproductive in Adam Smith's list. But we might also put aside at once the scene painters at the opera, the printers of trashy books, the makers of cloying sweets and noxious drinks. And if on the other hand we were called on to say what producers we should retain to the end, we should select not only those who supply the material things essential for existence—food, clothing, shelter—but also the physician who preserves our health and the teacher who maintains the education on which civilization

rests. The distinction between things essential and things dispensable is by no means the same as that between material and immaterial sources of utilities.

We conclude, then, that all those whose labors satisfy wants—all those who bring about satisfactions or utilities—are to be reckoned as taking part in production and are to be called productive laborers. Certain it is, whatever phraseology we care to apply, that no conclusions of importance for economics flow from the mere distinction between those who shape material wealth and those who bring about utilities of other kinds. And the test of the value of a distinction or classification is always that significant propositions can be laid down concerning the things put into a given class which do not hold for those outside the class.

This conclusion also enables us to dispose of an allied question: Is there nonmaterial wealth? Those who denied the old proposition—who maintained that labor which did not embody a utility in material objects was nevertheless productive—often maintained that there was such a thing as "nonmaterial" wealth. The phrase certainly is not in accord with common usage. We think ordinarily of wealth as something that can be kept and accumulated, and intend by it tangible things; and in this sense it is a contradiction in terms to speak of nonmaterial wealth. But if we use the more technical and therefore more precise phrase "economic goods," we include all those things and services which satisfy human wants and are not to be had free. The services of those whom Adam Smith and his followers called unproductive laborers come under this head. They are desired and prized, often highly prized; and they are yielded by human effort. The rewards earned by these efforts are an important topic in economic science, and the utilities provided are an important part of the sum of utilities which constitute, in the last analysis, the community's income. If we mean by wealth anything about which economic problems arise, we must make the terms coextensive with the term "economic goods"; and then we may speak of nonmaterial wealth.

§ 3. From this interpretation of the terms, it would seem to follow that all labor belongs to the productive class. If not only the

butcher and the baker are in this class, but the barber and the fiddler, do any remain who are to be regarded as unproductive?

Obviously, there are some persons who are outside the pale of productive activity. The paupers, thieves, swindlers, ne'er-do-wells, are parasites. Thieves and swindlers often exert themselves severely, tho not often continuously. But their activity is purely predatory. They contribute nothing; they simply try to get things away from others. Whether or no we should apply the term "labor" to their exertions, it is certainly not to be called productive labor.

A different question arises as to some labor carried on without violation of the law and without conscious delinquency, yet certainly of doubtful aspect. A quack medicine, containing ingredients which the maker knows to be noxious, or at best harmless, may be puffed by mendacious advertising into widespread use. Can it be said that the labor devoted to preparing it and persistently circulating lies about it is productive of satisfactions and therefore to be reckoned as productive labor?

To take another case, of still a different sort, what shall we say of the labor given in well-nigh all communities to the production and sale of intoxicating liquors? Among physiologists the settled conclusion is that tho the use of these stimulants in the lighter forms may lead to no harm, that of distilled spirits is in most cases bad. It is certain that an immense amount of misery and vice comes from the widespread use of strong liquors; that the diminution in their consumption during the last generation or two has brought betterment for mankind; that the world would be a much happier place if drunkenness could be stamped out. What has the economist to say of labor given to the production of things harmful?

These cases call for discrimination. They may be cases of fraud and deceit. They may be cases of wants misdirected, but none the less wants really felt and really satisfied.

Fraud and deceit mean that a person does not secure that which he expected and was led to expect. In an ordinary sale the seller is not presumed by the law to give a guarantee as to the quality of the thing sold: *caveat emptor*. But where a guarantee is given,

or a precise description equivalent to a guarantee, the buyer has a remedy in the courts.

The distinction made by the law is substantially that which the economist would make. The quack medicine may be a draft of flavored water or disguised alcohol. But so long as the purchaser wants this sort of thing, and buys because he has a notion it will do him good, the purveyor adds to the sum of satisfactions. The case is different where the purchaser wants one thing and is deceived into taking something else, since then his felt wants are not satisfied. Intermediate is the case where the purchaser does not know precisely what he wants and is wheedled into taking something which the other man wants to sell. Here it is often difficult to draw the line. Is the buyer foolish, or is he swindled? Does the seller lie outright, or is he merely expansive in praise of his wares? What the law can do is to aid in making the situation clear; and this is particularly needful where the consequences of misunderstanding are serious. Hence the pure-food and pure-drug legislation, and the legislation requiring that the composition of nostrums be precisely stated on their labels.

Where the want is really felt and really satisfied, the labor that brings satisfaction must be adjudged by the economist productive; and this even tho the ultimate consequences be harmful. The keeper of a dramshop is a productive laborer, even tho he purveys something which often causes misery. To enter on inquiries about the final effect on happiness would raise many questions of a different sort from those within the strict range of economics; inquiries which, if consistently followed in all cases, would range into almost every field of knowledge. There are intelligent persons who believe that meat, tho men like it, is unnecessary for nourishment and is frequently a cause of disease. Others maintain that such stimulants as tea and coffee are of ill effect; that health and happiness are promoted by abstinence from them. To judge between these various advocates and reformers is no part of the essential task of the economist. So long as a person who buys a thing or pays for a service really *desires* it, the labor which yields him the satisfaction is productive. Hence it has been suggested that we speak not of satisfactions or utilities,

but merely of "desirabilities"; a term that does come closer to what the economist has in mind when fixing the boundaries of his subject, but has the disadvantage of awkwardness. At all events, it is in this sense that the economist is concerned to inquire what labor is productive and what is not.

A case which may call for nice distinction between labor that is productive even tho morally questionable and labor that is predatory, is that of the professional gambler. The luxurious establishment at Monte Carlo may be regarded as simply purveying to that love of games of chance which is so universal as almost to be classed as inborn. So far as it does this—so far as the act of gaming is pleasurable to the customers—a desire is satisfied, even tho it may contribute to permanent welfare that such craving be kept in check. On the other hand, so far as both croupier and gamester are merely trying to get each other's money, and care not for the play in itself, the activities of both are predatory. Just what motive underlies the gamester's wagers may be a matter for nice psychological analysis. No doubt the two distinguishable motives—love of play and cupidity for the other man's money—are often combined. There are certainly instances enough where the pleasure of the play counts for nothing and then the keeper of the gambling establishment is simply predatory.

Returning now to such articles as were considered a moment ago—drugs and alcoholic spirits; whose effects may be noxious—we may note the obvious distinction between saying that a given kind of labor is productive and saying that it ought to be exercised. Tho a desire may be satisfied by the labor, it does not follow that happiness, or the best kind of happiness, is promoted thereby. The law may prohibit horse racing or gambling, or the manufacture and sale of liquor, because it is thought best that men should not have the gratifications at all. Whether prohibitions of this kind should be enacted raises questions, to repeat, of very wide range, to whose solution the economist can doubtless contribute but on which he says by no means the final word. The labor which yields a service may be in the eye of the economist strictly productive, but may be a kind of productive labor that had better not be exercised.

§4. The meaning which we affix to the word "productive" is further illustrated by one of those professions which Adam Smith regarded as indeed grave and important, but none the less unproductive—the law. With the lawyer may be grouped the judge, the policeman, the jailer. In a sense, their services are not necessary. They do not conduce directly to the production of material goods or to the rendering of services or utilities to consumers. They are adjuncts to the processes of production rather than immediately contributing factors. If all men were honest, truthful, fair-minded, willing to abide at once by the decision of an impartial arbitrator, the work of the legal profession and of all its hangers-on could be dispensed with, or at least reduced to insignificant dimensions. If virtue were universal, policemen and jailers would disappear, and lawyers would have little or nothing to do. Yet the experience of all peoples shows that—men being what they are—the work of the legal profession becomes indispensable in any complex society organized on the basis of private property. As property is accumulated and diversified, as exchanges between men multiply, as the precise relations between different persons come to be carefully defined by law, the business of interpreting the intricate system is put into the hands of a separate profession. The settlement of differences is intrusted to judges; the orderly conduct of affairs is aided by the advice of lawyers; the observance of the law is enforced by the police. No doubt an ill-devised legal system entails more labor of this sort than would suffice under a better system, and the unprejudiced observer may well question whether the law of our modern communities works as efficiently as it might. But a clumsy instrument, tho it involves more labor than one well adjusted, is none the less useful.

Similar considerations apply to the army and navy. The immediate object of the soldier's work is destruction. He must be supported by the rest of the community; he does not contribute directly to its well-being. Yet military protection has been thru almost all history an indispensable condition for the sustained conduct of peaceful industry. Like the policeman, the soldier is needed because of the bad passions of man. And even where de-

fense is not necessary, and armaments are maintained from national vanity or senseless rivalry, the soldier nevertheless must be reckoned productive in the sense that he does what people wish to have done and what they pay him for. The army and navy may be no more than dangerous playthings. But men are no less foolish when they pay for tawdry ornaments or vulgar amusement. It is not for the economist to sit in judgment on their tastes.

There is indeed a situation in which a military force is, from the economist's point of view, clearly unproductive. This is where it is used solely and simply for aggression. A pirate is obviously not a productive laborer. Unfortunately many of the heroes of history have been no better than pirates. The armies of the first Napoleon swarmed over Europe, levying tribute wherever they penetrated. No doubt deep-lying historical forces served to bring on the wars of the Napoleonic period; some conflict was inevitable between the old feudal order and that new order which arose with the French Revolution. But the domineering spirit of Napoleon turned the conflict in its later stages to mere aggression on the one side, exhausting defense against aggression on the other. That defense was necessary; yet all the effort applied both to offense and defense was in the last analysis a fruitless application of labor.

Lest this mode of considering the military be judged shallow by some of my fellow economists let it be added that the bare economic side of the matter is not the only one to be considered. Complex political and social questions present themselves, quite beyond the scope of a book on economics. No range of topics brings out more clearly the need of considering problems that are partly economic from other points of view as well. Even as a problem in economics alone, the industrial progress of mankind has often proceeded in strange ways. Civilization has gone forward on the powder cart, as in the War between the States (1861–65). Aggression itself sometimes leads to happier ends. The English first took possession of India in a spirit of sheer rapacity. Yet their rule has much promoted the material welfare of the native races. And in the conflicts between civilized peoples also, whatever their origin, a better order and a higher prosperity have emerged from wars that were rapacious. Reflections of this sort will occur

to every thoughtful reader, and lead him to qualify and interpret what has here been said of the relating of armaments and wars to the principle which underlies the conception of productive labor.

§ 5. There remain to be considered questions as to the relations of certain kinds of activity to the productiveness of labor. Are any of the business doings which go on in modern society to be judged unproductive?

When unscrupulous persons solicit funds from the gullible, ostensibly for "investment" or "speculation," and in due time run off with the money, their labor, systematic and strenuous tho it may be, is obviously predatory. Not only they but the clerks and assistants whom they employ (whether these be accomplices or innocent) are unproductive. Now it is alleged that, outside the range of operations so clearly predatory as to be made criminal by law, there are others, within the pale of the law, whose economic effect is substantially the same. This is said, to take a familiar example, of speculative transactions in general. In our highly organized modern communities an immense amount of buying and selling is done for a turn in the market. A man buys wheat or cotton which he does not want and which never gets into his possession; he promptly sells his nominal title at an advance in price, pocketing what is called a profit. Is any contribution made to the sum of utilities by such transactions? It may be assumed that the pleasure of the game, which may be an element in gambling with cards or dice, here plays but a negligible part; the motive is simply to get gain somehow. The most conspicuous operations of the sort are on the stock exchange, where sales and purchases take place on an enormous scale with no traceable effect in contributing to production or to social income. The business involves an elaborate apparatus—brokers, clerks, officers, a periodical press of its own. As the clerks of a bare swindler are unproductive, so must be those of the broker, if he is himself in the parasitic class.

But this sort of accusation has been pushed further. A large part of what is ordinarily called "business" has been placed under the same ban. Not only those who are usually called speculators but those who "operate" in real estate—buy and sell land for a margin

of profit—and the bankers who "handle" stocks and bonds are described as mere parasites. Nay, all business men of every kind have been condemned by socialist writers as essentially unproductive—that is, so far as they are not directly doing work of management and superintendence. By the socialists "business" has been adjudged simply a way of securing a gain thru the ignorance or weakness of others, and therefore to be condemned as useless to society.

The questions here raised cannot be answered until after a consideration of some very complex matters. But the mode in which they should be dealt with and the nature of the answers to be sought can be indicated now, even tho with some anticipation of later conclusions. Thus, as regards one of the set of operations supposed to be unproductive—speculative dealings—it must be admitted that the charge is in part founded. Tho some speculative dealings in commodities and securities serve a useful purpose, others are in large part mere wagers, akin in their economic effect to ordinary gambling.[1] Judged by the test which we have set up—whether the labor adds to the sum of satisfactions—all those who engage in mere wagering speculation are unproductive laborers: not only the principals, but the brokers who execute their orders, the clerks who record them, the mechanics who put together and operate the "tickers" in the broker's quarters. All belong in the class whose work serves no useful end.

The same test is to be applied to the activity of business men, but here the balance of gain is much clearer. Tho the greater part of speculative dealings is probably of no utility, the greater part of business men's doings has great utility. The indictment of those socialists who charge that they are predominantly unproductive overshoots the mark. The function of the manager or leader of industry is of high service in production; and this even tho, like the banker, he may merely advise and select and promote, taking no direct part in the management of industry. He adds, often adds much, to the abundance of commodities and the satisfaction of desires. But it is none the less true that in any large center of industry there will be found plenty of persons engaged

[1] Compare Chapter 11.

in "business" whose doings are essentially parasitic. They pick up a living, perhaps a very comfortable one, by shreds and patches of dealings, by shrewd or fortunate guesses in buying and selling, by waiting for land or securities to rise in value. Often they are sober, solid citizens, personally estimable; so indeed as a rule are the stockbrokers who provide the facilities for the gambling speculators. These respectable persons would resent with indignation the suggestion that they belong in the predatory and parasitic class. But one of the most remarkable phenomena presented to the student of economics is the ignorance of all sorts of persons regarding their place and function in the industrial world. The broker or merchant, no less than the mechanic or clerk, sees the little corner in which he is at work, and knows nothing of its relations to the community as a whole. The respectability of an employment and even the spirit in which it is pursued give no certain clew to its effect on the general welfare.

It is the aim of the legal system under which we live—the system of private property—to inhibit predatory doings. Hence not only physical violence, but fraud and deceit, are forbidden and punished. This aim of the law is in the main attained. He who earns his living in a lawful manner commonly contributes to the sum total of satisfactions. The view, sanctioned more or less explicitly by some socialist writers, according to which the work of manual laborers alone is productive and all the income-earning and money-making of the well-to-do classes are unproductive, carries the indictment against the existing system too far. But the fact that criticism against the working of private property is exaggerated should not blind us to the fact that there exist opportunities for securing an income or even amassing a fortune, not beyond the pale of the law, yet of a kind which the economist must regard as predatory, and so unproductive.

Some opportunities of this kind are due to imperfections in the law as it stands. With changes in economic conditions, proceedings that once seemed helpful to the promotion of the general welfare, and perhaps at one stage were helpful, cease to be so or remain so only in part. Thus joint stock companies, or corporations, have proved a device of great efficacy in furthering improvements in

the arts and in securing more abundant and varied production. On the other hand, the statutes under which corporations may be organized, especially in our American states, have often made possible mere thimblerigging and plundering. The reform of the laws of incorporation in such a manner as to keep the good and reject the evil is still one of the pressing problems in the United States.

To discriminate clearly between the operations that are in the end helpful toward satisfying desires and those that are not, is sometimes impossible even after the nicest weighing of the results by the best judges. The law, for instance, withholds its sanction from mere wagering contracts. Yet transactions which are wagers cannot be distinguished in outward form from others which are useful to society. There is a vague consciousness in the public mind that some persons are engaged in "legitimate" business, while others doing the same sort of thing are "plungers," occupied "illegitimately." But to draw a precise line between those that may be approved and those that may not, is no less difficult for the business man, however intelligent and wide-minded, than for the judge or the economist. So it is with the law of fraud and deceit. As long as men are free to choose for themselves and act according to their own judgments, those who are shrewd and watchful will make better bargains than those who are dull and unobservant. When does one man overreach another, when does he simply leave the other to judge for himself as to his own interests? The probabilities are that for the sake of securing the general benefits that flow from private property and competitive dealings we shall always have to permit some doings that are on the line between the productive and the predatory. If the law brings it about that labor is applied in the main to the satisfaction of desires; if most of the unproductive doings are prevented; if the system as a whole works well and the predatory operations are only its loose ends—it will be better to accept some evils as inevitable and to set off against them the general benefits.

CHAPTER 3

THE DIVISION OF LABOR AND MODERN INDUSTRY

§ 1. Two forms of the division of labor: the simpler and the more complex.—§ 2. Advantages from the simpler form: dexterity, continuity, adaptation to aptitudes.—§ 3. Advantage from the more complex form: the development of machinery. The industrial revolution of the eighteenth century. The use of nature's power.—§ 4. Division of labor means unconscious cooperation. Exchange.—§ 5. Exchange formerly covered a limited economic area. Cheap transportation enlarges the area; railways.—§ 6. Wider markets bring more minute division of labor. Illustration from butcher's trade.—§ 7. The geographical division of labor, illustrated by the United States and Great Britain.—§ 8. Two sorts of gain from the geographical division of labor.

§ 1. THE division of labor is one of the great central facts in modern society. From this arise some of the most difficult questions of economic theory, the most common popular fallacies, the most serious problems of legislation.

The division of labor may be analyzed under two heads. On the one hand there is the simpler form, under which a workman completes the whole of one of the stages in production; the tailor, the cobbler, the carpenter, ply their several trades. On the other hand there is the more complex form, under which there is a splitting up of several operations all belonging to one connected process of production. In more primitive stages of industry the shoemaker might be a tanner, and the whole process of converting the rawhide into a shoe thus be in one hand. Nowadays, the shoe itself is not put together by the cobbler; it is the work of a large number of different workmen in a factory, of whom some do nothing but cut the leather, others stitch it, others put on the soles, still others the heels, and so on, with an elaborated parceling of different operations.

Obviously, a hard-and-fast line cannot be drawn between these two forms. No craftsman carries thru from beginning to end any one operation in production. The tailor buys his materials from

the cloth maker; the cloth maker buys his wool from the farmer or grazier. The cloth maker and the grazier in turn buy the tools fashioned by the mechanic, who buys materials from the ironworker and woodworker. On the other hand, the tailor does not necessarily carry his own work thru even the whole of the stage with which he is concerned. It may be divided between the cutter and the stitcher; and similarly the cloth maker's work may be parceled out between the weaver, the fuller, the dyer. The difference between the simpler and the more complex division of labor is essentially one of degree. Nevertheless, this difference of degree is important. The two sorts of arrangement bring about somewhat different advantages and give rise to different social conditions.

§ 2. Let us consider first the simpler division of labor. This dates far back into antiquity. The familiar crafts are of very old standing. The extent to which their names have been adopted as surnames shows how occupations were separated in a comparatively simple state of society, such as that of the Middle Ages, when patronymics were in process of formation. The Carpenters, Masons, Smiths, Weavers, Drapers, Tailors, Dyers, Saddlers, Shoemakers, Millers, Bakers, Coopers, and such other common surnames indicate what sort of division of labor was maintained with comparatively little change for hundreds of years.

The chief advantage in production from this form of the division of labor is the gain in dexterity which comes from the constant practice of the same occupation. So familiar are we with the effect of practice that we assume as a matter of course the skill which comes from it. Reading, writing, the donning of our clothes and the lacing of our boots are effected with ease, almost without effort, from the ingrained effects of habit and iteration. Piano playing and typewriting are marvelous to the inhabituated, easy to the point of indifference for the practiced hand. The acquired dexterity of the craftsman and mechanic make their productive capacity vastly greater than it would be if each had to carry on a dozen occupations and were at best half proficient in any one.

Other gains have also been enumerated as accruing from the simpler division of labor. There is a saving in time when the same task is followed without interruption. The carpenter, even

tho no more quick and ready than the farmer, can yet accomplish more in the hour or the day than the farmer who tries to do jobs of tinkering in his spare moments. Something also is due to the adaptation of tasks to the abilities of the workers. There are differences between the inborn abilities of individuals even as regards tasks for which training and practice are the more important causes of dexterity. Among mechanics a certain proportion only have the sure eye and the deft hand which are required for the most exacting tasks. It is obviously advantageous that they should confine themselves chiefly to these, leaving the less exacting to persons of ordinary capacity. Even for comparatively simple occupations there are differences in the qualifications of individual workmen. The work of a motorman on an electric car seems of the most monotonous sort, easily accomplished by any adult. Yet it requires a certain steadiness and alertness of attention not possessed by all laborers. How far differences of this sort are the result solely of inborn qualities, how far brought about or accentuated by education and environment, need not here be considered. So long as they exist, there is a gain if each individual is called on to do only that for which he has the greatest aptitude.

The last-mentioned factor in the division of labor—the adaptation of tasks to varying aptitudes—is of most importance as between those who work with their heads and those who work with their hands. Tho mental training has its marked effect as well as manual training, and tho instruction and practice tell in the lawyer's trade as well as in the mechanic's, inborn abilities are important in greater degree for the former. This is more particularly the case in all work which calls for initiative, superintendence, direction. There is a difference of far-reaching consequence between those who have the qualities for leadership, whether in the arts or in intellectual life, and those who must belong to the rank and file. A very great gain is often achieved when those who are born leaders can devote themselves solely to the work which they alone can do, or which they can do best, leaving to others with no such capacities the routine mechanical or clerical work.

The great mass of men, however, have no special aptitudes. For them, continued practice, begun or aided by systematic

training, is the chief cause of skill in any particular sort of work. In the main, the division of labor is a cause rather than a result of specialized capacity. Most dextrous men are so because they have long practiced an art; they do not practice it because they are born with dexterity.

§ 3. Let us turn now to what we have styled the more complex form of the division of labor. This is the salient characteristic of the development of industry during the last century and a half; a development which has gone on with accelerating pace in very recent times. The change in industry and the nature of the new order of things can be described most concisely by saying that the tool has been replaced by the machine.

Tho the gain in efficiency from the division of labor arises chiefly from the skill acquired by repetition, none of the trades familiar under the simpler division of labor was reduced to the continuous repetition of identical movements. The carpenter, the mason, the smith, the tailor—each was master of his trade as a whole. While gaining proficiency from unceasing practice, he yet turned from one part of the occupation to another. The instruments which these artisans used were tools of varied kinds, adapted to the different parts of their occupations. A "tool," as that word is still commonly used, means a hand tool, put in motion by human force and requiring adaptation, judgment, flexibility.

The gradual elaboration of the division of labor slowly enlarged the number of occupations, diminished the range of each one, and tended to reduce each more and more to an identical routine. Thus the making of cloth was divided between the spinner, the weaver, the fuller, the dyer. When the steady repetition of the same movement becomes an important part of an industrial art, it is possible to apply other force than that of man's muscles. It is true that no machine, even in the highly elaborated forms of modern times, can rival in dexterity and flexibility the human hand. But whenever the same thing is to be done over and over, the blind forces of nature, working thru a machine, can do it as well as any human hand, and indeed better than most human hands. The division of labor in its simpler form gradually was developed to the point where the application of power was

possible. The gain from the application of power proved so great that there was a reaction on the division of labor, a motive to split up the steps in production still further, to reduce more and more of them to the repetition of identical movements and so to make possible in still greater degree the use of natural forces.

The great change set in markedly during the second half of the eighteenth century. The division between the spinner and the weaver, one of the oldest elaborations of the division of labor, gave occasion for an epoch-making application of machinery and power. In 1764, Hargreaves invented the spinning jenny; in 1769, Arkwright brought out his rival spinning machine; in 1779, Crompton invented an apparatus which combined the devices of Hargreaves and Arkwright, and brought the spinning machine to a still further stage of perfection. All three were directed to the mechanical repetition of the twisting of the fiber, and water power was soon applied to setting them in motion. Not long afterwards, weaving was also subjected to the same principles. The power loom was gradually elaborated, and in the beginning of the nineteenth century began to supplant steadily the hand loom. By the close of that century, the old-fashioned weaver's trade had become in advanced countries like England and the United States, a thing of the past. The textile material to which these inventions were first applied was cotton, for this has an even and homogeneous fiber which makes it most readily available for machinery operated continuously at uniform speed. Wool, linen, and silk, being of less even fiber, were subjected to the machine process later than cotton, thru a long series of subsidiary inventions. It was not until late in the nineteenth century that silk, the most delicate and irregular of these fibers, came to be manipulated on a large scale by power machinery.

Water power was used for the textile manufactures in their earlier stages; but it was soon supplemented and largely replaced by the steam engine. The steam engine was brought by Watt to the stage of effective working in 1781. It was first used on a large scale for the pumping of water out of mines—an obvious case for the application of power, since it calls for the unchanging performance of the simplest movements. It was soon applied further,

not only to the textile industries and to a wide range of other manufactures, but to transportation. Steam was used in navigation by Fulton on the Hudson River in 1807. An even more important application of steam to transportation came when the locomotive was perfected by Stephenson in 1830. This created the modern railroad and, as we shall presently see, marked the beginning of a still further development of the division of labor.

The series of great inventions of which these were the most conspicuous, brought about what is known as the Industrial Revolution—a change in the arts, and a consequent change in economic and social conditions, greater than has appeared during a like short time in any stage of human history. This in a way states the case too strongly. The change, while revolutionary, was not sudden or fortuitous. A long series of inventions and devices, reaching far back into the centuries preceding, prepared the way. Yet a rapid burst and flowering there was in the second half of the eighteenth century. Its fundamental economic characteristic was an elaboration of the division of labor, thru the splitting up of the stages of production into separate operations, each one of which is repeated continuously and so may be carried on by the machine. The carpenter's sawing, planing, joining, molding—each of these is now done separately by machinery, usually in establishments that tend steadily to become larger and larger and to subdivide still more the various operations of the trade. The cobbler of former days put together a shoe for himself; in a modern factory the shoe goes thru some eighty different processes. In the manufacture of files there were (1912) in one large establishment ninety separate operations or handlings thru which each individual piece had to go from the time the steel reached the factory until the time when the finished files were ready for sale. So it is with ironworking, with all the elaborated processes of the textile industries, with printing and book-making, not least with the very making of machines and tools. The machines now used are vastly more complex and more efficient than was dreamed of in the early stages of the application of power, and have extended the principle of the automatic repetition of identical movements to tasks long thought far beyond its scope. The work of the hand is not indeed

superseded; the skillfull workman and the adaptable tool retain a large place in industry; but the range of their work tends to become more and more restricted. Within each branch of industry, as one stage after another is subjected to the machine process, the remaining stages have a narrower and simpler range, in which inventive spirit constantly finds new opportunities for the application of power. Thus the character and the working of the division of labor have been profoundly and all but universally modified.

The essential gain from this modern development of the division of labor has come from the virtually unlimited store of natural power. Once identity of movement has been secured, there is no work so heavy, no operation so delicate, but that the machine can repeat it day in, day out. Human labor, applied first to putting together the machine, then to guiding the natural forces that move it, accomplishes vastly more than the same amount of labor applied to the making and using of the simpler tools of former days. Coal and falling water are the great sources of power; and tho nature does not supply them without limit, the application of machinery has not yet been fettered for human needs by any limitation, nor is it likely to be fettered in the future as far as we can look forward into it. The labor required for any one operation in production has been immensely lessened by the industrial changes of the eighteenth and nineteenth centuries, and appears likely to be lessened no less rapidly and largely in the twentieth century.

The period in which we live has been aptly called the age of machinery. Its characteristic phenomena are mainly the results of the use of machinery, and they will engage our attention in many parts of our subject. They are seen in the growth of capital, and the increasing power and importance of the business man who has control of capital; in the spread of production on a large scale, and the tendency toward monopoly in many branches of industry; in a new position of workmen, a wider gap between employers and employees, and a consequent development both of labor organizations and of employers' associations; in grave social problems from the employment of women and children in factories; not least, in a loss of individuality in the working population and a strengthening of the lines of demarcation between social

classes. Of all these consequences of the complex division of labor more will be said as we proceed.

§ 4. The division of labor obviously means that the persons who carry on the several operations of a given branch of industry combine to bring about the final result. It means, no less clearly, that those engaged in different industries combine to satisfy the varied wants of the community. Each contributes his special product to be used by all; each uses the products contributed by the rest. The division of labor may thus be described also as the combination or coöperation of labor.

Combination of labor may conceivably be carried out deliberately, with conscious control and coördination, with immediate sharing of the joint output, and without exchange. In the ancient civilizations of Greece and of Rome we get glimpses of establishments of the rich and privileged in which the several trades are plied by slaves for the benefit of the whole household. In the earlier Middle Ages also we find seigniorial possessions, where the serfs have specialized occupations and contribute in kind to the lord's needs. Even in modern times, we have had examples of communistic societies, in which there is a division of labor among the individual members, yet no exchange, each member contributing his part to the common income and each receiving from that income a share deemed equitable. Such a society does not approach so nearly to self-sufficiency as the ancient household or the medieval estate; it must buy and sell on a considerable scale with the outside world, whereas those earlier organizations bought very few things (salt and iron, for example). Yet within its own limits the division of labor leads to no exchange between members.

Commonly, however, the division of labor has brought with it as a natural corollary the *exchange* of the several commodities produced by different workers. The cases noted in the preceding paragraph are comparatively rare in economic history; at all events, they give no clew to the phenomena of the modern industrial world. In this the division of labor almost always means exchange, and the relation between the workers is very different from that in the community where there is conscious and deliberate combination of effort. It is strictly true that the workers in a

modern society combine in bringing about a joint output; but the consciousness of coöperation is lost. The individual is not thinking of the joint output. Only if he happens to be versed in the books and theories of economic writers, and bears them in mind in his active hours, is he aware that he is performing one small operation toward a joint output and is sharing in the manifold contribution which others make to that joint output. The things on which he works are not part of a common store, but are private property, bought and sold, cared for and guarded, by each individual for himself. He thinks only of the particular product which he sells, and of the terms on which he can buy other products. He is intent on the results of the exchange thus made, and tries to secure for himself the best terms of exchange. The phenomena of exchange are the dominant phenomena of the modern world.

§ 5. For some centuries preceding the industrial revolution of the eighteenth century, the typical form of exchange was that between the small city or town and the agricultural region immediately surrounding it. This was the period of the simpler form of the division of labor, of the familiar handicraft; the period of the tool preceding the modern period of the machine. The city of early modern times was the center of an industrial community which was in the main self-contained. Within the city the burghers carried on the arts and crafts, to it the surrounding rural population brought food and materials, and in it they made their purchases. The city craftsmen were united in the gilds which were so conspicuous a feature of the economic organization of that period. Each craft was open only to the members of a gild, who trained apprentices and employed journeymen, and transmitted from generation to generation the knowledge of its trade. The organization of the gilds, and the regulation and restriction of their membership, were inevitable and doubtless beneficial at the outset, assuring protection and mutual aid, and the maintenance of skill in the arts. In later times, their regulations were made the means of monopoly; they had long outlived their usefulness even before the great inventions of the industrial revolution put an end to the economic organization of which they were a part—aspects of the gild system not closely related to our present topic. So far as it

bears on the division of labor, it was part of what the Germans called *Stadtwirthschaft*—the city organization of industry. A map of England and of the greater part of western Europe from say 1350 to 1800 is dotted with a large number of cities of modest size, each the center of a more or less isolated economic area. Tho there was some exchange of special commodities between different countries and between the different economic areas within a country, the main exchanges were between the city and the surrounding agricultural district, and the characteristic stage of the mechanical arts was that of the simpler division of labor between the familiar crafts organized in the medieval gilds.

The steps through which this organization of industry has been replaced by that characteristic of modern times were at first slow and gradual. But in the eighteenth century the industrial revolution brought a sudden burst of great changes. Without pausing to consider the events of the sixteenth and seventeenth centuries, which prepared the way for these changes, we may contrast the final result with the conditions of the early simpler division of labor, and so understand better the conditions of our own day.

The economic area has been immensely widened. It has come to include the whole of a country, in some respects the whole of the world. There is division of labor not only between the different crafts within a city but quite as much between different cities and countries. On the other hand, the crafts themselves have been split up into more minute subdivisions, and different parts of each are practiced in widely separated localities. These tendencies have been immensely promoted by the modern improvements in transportation—improvements which have themselves been the results of the introduction of machinery. The use of power, especially thru the steam engine, was the dominant factor in the industrial revolution; and in no direction has it had larger effect than from its application to traction and to navigation.

An epoch-making invention was that of the locomotive. Roads had been much improved in England during the latter part of the eighteenth century, when Telford and Macadam devised their methods of constructing roadways. During the same period canals had also been dug and used to no small extent both in France and

in England; and the people of the United States, always driven by their special industrial conditions to search eagerly for improvements in transportation, pushed the use of roadways and canals in the first quarter of the nineteenth century. But in 1830 came the locomotive. In this case, as in that of the steam engine and indeed of almost all great advances in the arts, the final attainment of the successful device was due to a long series of experiments by many contrivers. Stephenson in 1830 perfected rather than invented the locomotive. So the modern railway was created; and thereby began a second industrial revolution, or at least a second phase of the industrial revolution. Side by side with the railway have been the great improvements in water transportation. The application of steam to navigation, thru the paddle wheel, was a comparatively simple matter and was accomplished early in the nineteenth century. But the paddle-wheel steamer was too clumsy, too liable to damage in storm, for communication across great bodies of water. It was not until Ericsson's invention of the screw, in the middle of the nineteenth century, that ocean navigation underwent a great change. This change in any case was not so far-reaching as that wrought by the railway; for water transportation by sailing vessels had always been comparatively cheap whereas land transportation had been slow and dear, and its dearness had imposed great obstacles to the division of labor within large land areas.

§ 6. As Adam Smith remarked in 1776, in the earlier stages of the modern era, the division of labor is limited by the extent of the market. The village cobbler will turn out no more shoes than it is possible to dispose of within the economic area he can reach. To divide the work of shoemaking between the cutter, the stitcher, the heeler, the laster, is not feasible unless as many shoes can be marketed as the combined labor of all will produce. A modern shoe factory, with its elaborate machinery and highly developed division of labor, turns out thousands of pairs of shoes daily. These shoes can find their purchasers only when a large population can be reached from the central source of supply.

Many other illustrations could be given of the way in which the division of labor has been pushed farther and farther with the extension of the market consequent on cheapened transportation.

Furniture is made nowadays in large factories, often placed near the sources of timber supply and distant from the persons who are to use the articles. The cabinetmaker of olden days has been replaced by workmen who tend and direct a series of machines, each of which performs unceasingly its part in the operations of sawing, planing, grooving, turning, polishing. Plows are no longer made by the village blacksmith, but put together in a great factory and then distributed over the earth. Unless it were possible so to distribute them, plows could not be made in quantities at the factory, and there could be no elaborated division of labor in making them. One of the most striking results of the widening of the market is seen in the transformation of the butcher's trade. Until the latter part of the nineteenth century, the butcher carried on his work as he had done it for thousands of years before. His cattle came from near-by farmers, and the meat was supplied to near-by customers. Thru the larger part of the United States he has now been supplanted by the great packing establishment, where cattle are slaughtered by the thousand. In these establishments dozens of different stages in dissecting the carcass are allotted to as many different sets of workmen. The application of power has not here been carried as far as in some other industries; yet at every stage where it is possible, the machine is set to work; and even where it is not, the workman is assigned to the unceasing repetition of a single operation.[1] Every part of the animal is used, and every part

[1] "It would be difficult to find another industry where division of labor has been so ingeniously and microscopically worked out. The animal has been surveyed and laid off like a map; and the men have been classified in over thirty specialties and twenty rates of pay, from 16 cents to 50 cents an hour. The 50-cent man is restricted to using the knife on the most delicate parts of the hide (floorman) or to using the ax in splitting the backbone (splitter); and, wherever a less skilled man can be slipped in at 18 cents, 18½ cents, 20 cents, 21 cents, 22½ cents, 24 cents, 25 cents, and so on, a place is made for him, and an occupation mapped out. In working on the hide alone there are nine positions, at eight different rates of pay. A 20-cent man pulls off the tail, a 22½-cent man pounds off another part where good leather is not found, and the knife of the 40-cent man cuts a different texture and has a different 'feel' from that of the 50-cent man. Skill has become specialized to fit the anatomy. . . .

"The division of labor grew with the industry, following the introduction of the refrigerator car and the marketing of dressed beef, in the decade of the seventies. Before the market was widened by these revolutionizing inventions, the killing gangs were small, since only the local demands were supplied. But when the number of cattle to be killed each day increased to a thousand or more, an increasing gang or

is manipulated on a large scale under a further minute division of labor. The output in its varied forms—the meat of all qualities, the fat, the hide, the bones, the horns, the very hair—all is then marketed to millions of people, distant hundreds of miles, sometimes thousands of miles, from the packing establishment. All such elaborate organization and division rest on the possibility of transporting the products a great distance, and so supplying an enormous population from one central point.

§ 7. The great improvements in transportation during the nineteenth century have given immensely wider scope to a phase of the division of labor which we have not yet considered. This is the geographical division of labor.

In medieval and early modern times, those articles only could be transported for any considerable distance which had great value in small bulk. Such were drugs, spices, fine cloths, rare silks and cottons, choice weapons and armor. These were used chiefly by the small circle of the rich; trade in them did not affect the mass of the population. Where water transportation could be used, there was indeed some possibility of trade and exchange in the bulkier commodities. For this reason England, with her insular position and much-indented seacoast, was able at a comparatively early stage to export such commodities as wool, copper, and tin, and to develop in some degree the geographical division of labor. With the improvement and enlargement of vessels, the greater security of the seas, and the use of the mariner's compass, trade by water gradually grew to greater and greater dimensions. A still further extension came in the latter part of the eighteenth century, when parts of the interior of the civilized countries were tapped by canals. The most far-reaching development of the geographical division of labor came with the railway; for the railway can reach all parts of the land. The industry of almost every part of the world was transformed by this mighty solvent.

The United States at the present time presents what is probably

crew of men was put together; and the best men were kept at the most exacting work."—Professor J. R. Commons, in the *Quarterly Journal of Economics*, Vol. XIX, pp. 3, 6. It will be noticed that here there seems to be scope for that advantage from the division of labor which is secured from the adaptation of the tasks to the varying abilities of the several workers. Cp. p. 31, above.

the most extreme case of geographical division of labor highly developed under the influence of cheap transportation. The southern part of New England is a manufacturing hive. The food and raw materials there used come from all parts of the world. The wheat and other breadstuffs come from the Mississippi and Missouri valleys; the meat and animal products from the same regions, and some from regions farther west; the cotton from the southern states; the wool from the trans-Missouri region, as well as from Australia, Argentina, China, Siberia. All sorts of manufactured articles are sent from New England in exchange—cotton and woolen fabrics, boots and shoes, metal wares, tools and machinery. The anthracite district of eastern Pennsylvania, again, is given wholly to the mining of hard coal; all its manifold supplies come from without. Pittsburgh is the center of a district in western Pennsylvania given mainly to the mining of bituminous coal and to manufactures which use that fuel, such as iron and steel and glass. Here too the food, clothing, articles of comfort and luxury are obtained from all parts of the United States and of the world. No part of the country is self-sufficing; each is constantly sending its products to distant regions, in return receiving the products of distant regions.

An example no less striking of the geographical division of labor is to be found in the condition of Great Britain. That country imports the greater part of its food—four-fifths of its breadstuffs and more than half of its meat and other food supplies. Its wheat comes chiefly from Canada, the United States, Russia, Argentina; its meat very largely from the United States and Australasia. All the cotton and almost all the wool which serve to clothe its people are brought from other countries. These various commodities, as well as the others which come from tropical regions, are obtained in exchange for a great range of manufactures exported. The people of Great Britain, by devoting their labor chiefly to manufactures and exchanging them for the imported foodstuffs and raw materials, get vastly larger returns than they could by producing everything at home. New England and old England are in substantially the same industrial position. It is probable that neither could support its present population on its own soil; it is certain

that neither could satisfy in this way the imperative needs for food, clothing, shelter, warmth, except on very much harder terms and with very much scantier results. Each is dependent on trade with other regions, the main difference being that in the one case virtually the whole of the trade crosses the political border and in the other the larger part of it is within the same nation.

In consequence of this highly developed division of labor, the position of cities is essentially different from what it was in medieval times. They are no longer dependent for food and materials on the agricultural regions surrounding them, nor do these regions depend on the adjacent cities for their supplies of manufactured commodities. As regards the country surrounding them, the cities are centers for the distribution of goods rather than for production. Many cities have special articles of manufacture and in this sense are producing centers, but their specialties are disposed of over all the world thru the distributing centers. The very large cities, with a wide range of miscellaneous manufactures and with a great distributing trade, overlap in their economic areas.

The far-reaching complexity of this situation makes it extraordinarily difficult to keep the division of labor in working order when there is change in the underlying conditions on which it depends. In times of economic disorder and depression (about which much more will have to be said [1]) the intricate mechanism gets out of gear. The eventual buyers of the finished products—the consumers—prove unable or unwilling to buy what the distant chain of producers are equipped to supply and indeed are committed to supply. Depression, inability to sell what is produced, unemployment are the familiar phenomena of hard times. By way of remedy it has sometimes been proposed—the idea cropped out frequently in the great depression of the thirties of the present century—that a return should be made to the simpler conditions of earlier days. The city-dwellers, it was urged, should exchange their wares directly for the food produced in the surrounding country. But when once the division of labor is highly developed, this return to the simpler ways of earlier centuries is impossible. A shoe city like Brockton, a textile center like Lawrence, a me-

[1] See below, Vol. II, Chapter 43.

tropolis of varied manufactures like Philadelphia—each produces more of goods than its immediate agricultural region can use; the goods are made in quantities to satisfy millions of widely scattered consumers. And conversely the great agricultural regions grow vastly more of their products than the cities in their immediate neighborhood consume; these cities act mainly as traders or middlemen for the distant manufacturers. The country adjoining a large industrial city can use but a fraction of what the city is equipped to turn out, and the city in an agricultural region can use but a fraction of what the farmers habitually raise. There is always of course some direct exchange; some milk, some vegetables, some poultry from the near-by farms, and some purchases by them of the goods made in the city. Even this limited range has tended to become less with the perfecting of transportation and storage for even the most perishable of farm products, and with the surprising accessibility of the goods bought by the farmers. No rearrangement of this established organization is possible in short order; and if for any reason change in it there must be, the process is difficult and slow.

§ 8. The gains from the geographical division of labor are of two sorts, analogous to the two sorts of gain from the division of labor between individuals. In part they arise from the adaptation of different regions to the production of specific articles, and in part from the proficiency which is the result of exclusive application to one task.

The division of labor between tropical and temperate countries obviously brings the gain arising from specific adaptation. Tropical fruits, spices, coffee, sugar, are exchanged for the wheat and corn of temperate climes. The southern part of the United States, again, has a climate peculiarly adapted for growing cotton, while in the great central plains there is a corn belt and a wheat belt—great stretches of country with climate and soil peculiarly adapted to one or the other of the staple cereals. The abundant deposits of excellent coal in the western part of Pennsylvania cause that district to devote itself to coal mining, and to the industries for which cheap fuel is essential. Extraordinary deposits of iron ore have been found on the shores of Lake Superior and thousands of

workmen there mine the ore, procuring from other parts of the country all the varied articles which they consume. Italy has a climate adapted to the culture of the grape and of citrous fruits, and she exports them to the countries of more rigorous climate. Italy has no coal; she imports it, chiefly from the great beds of Great Britain. The enumeration might be indefinitely extended. It is obvious that there is a gain when the wheat and corn are produced in the regions favoring them, the iron and coal where they are most abundant, the cotton where the soil is best. The geographical division of labor is not indeed all-embracing; there are obstacles to its sweeping application from such causes as the force of habit and the cost of transportation. Yet there is a strong and steady tendency toward the pursuit of a branch of production in that place for which the natural advantages are greatest.

Different in origin and basis, tho the same in effect, is that division of labor between different regions which rests on the mere fact of specialization and acquired skill. Exchange between individuals, tho based in part on differences in native aptitudes, rests in the main on acquired dexterity. So it is in considerable degree between different regions. When once an industry is conducted on a large scale, with elaborate machinery and with a great output, it will tend to be concentrated. But there may be no strong reason for its concentration at one place rather than another. There is no cause in the natural conditions why Bridgeport and New Haven in Connecticut should be specialized centers for the manufacture of metal wares, Brockton in Massachusetts for shoes, Cohoes in New York for knit goods, Nottingham and Bradford in England for laces and woolen stuffs, Lyons for silks, Chemnitz in Saxony for hosiery.

For certain sorts of industries there is a gain from the mere fact that a number of establishments carrying on operations of the same sort are clustered together. Subsidiary industries spring up, supplying them with materials or accessories. When workmen skilled in particular operations are required, their selection and training are easier. The mere attractiveness which a city has for most persons makes it easier to secure and retain a steady force of laborers. Sometimes the first cause of the location of an industry in a par-

ticular place has been the energy, ingenuity, resource, of some individual. His capacity as leader builds up an establishment; others then follow his lead. Sometimes the natural adaptation of a spot causes an industry to spring up in that spot, and later to persist from the mere effect of acquired advantage. Thus some of the manufacturing cities of New England, such as Lowell and Lawrence, grew up on sites having water power, before steam power was as fully developed as in later times, and when the transportation of coal was more costly. It is doubtful whether the water power would now cause these centers of population to be built up; but being there, they tend to remain. All thru the broad, flat country of the Mississippi Valley there have sprung up cities and towns, of which one is the seat of the manufacture of vehicles, a second of furniture, a third of engines and machines, with no obvious causes why one place rather than another should possess the particular industry. In whatever place the industry is, the advantages of concentration are secured. A wide market from cheap transportation makes possible the conduct of the industry on a large scale and so the use of much capital, of elaborate machinery, of specialization, and minute division of labor.

A considerable part of the division of labor between nations and a large volume of the trade between them seem to rest on this second cause. Especially as regards manufactured articles, some countries have advantages in production which rest not on natural resources but on acquired efficiency. England's manufacture of certain kinds of woolen goods, the silk manufacture in France, perhaps the linen manufacture of the north of Ireland, present cases of this kind. This is the real basis of the argument for protection to young industries. So far as the division of labor between countries and the trade between them are the results of natural differences, they are best left to work out their results without restriction. But so far as they rest on acquired skill, there is at least a possibility that they may be superseded to advantage by similar division of labor and similar trade within the country.[1]

[1] See what is said on this subject in Chapter 37, § 2.

CHAPTER 4

LARGE-SCALE PRODUCTION

§ 1. Growth of large-scale production closely connected with the Industrial Revolution. Advantages of large-scale production: use of machinery, saving in general expenses, buying and selling, utilization of by-products, experimenting.—§ 2. Limitations, chiefly from difficulties of superintendence. The case of agriculture. Other industries. The development in retail trade.—§ 3. Scarcity of able managers as a cause of limitation. This human factor usually ignored by the socialists.—§ 4. Large-scale management. Combination, horizontal and vertical. The Steel Corporation as an example. Other examples. The tendency to vertical combination less strong than that to horizontal.—§ 5. Competition often wasteful; tho the waste is less than it seems.

§ 1. THE tendency to large-scale production has shown itself in all civilized countries since the industrial revolution. It has profoundly modified social as well as economic conditions, and bids fair to modify them still further in the future.

The characteristic features of the tendency are that the size of the individual establishment becomes larger and the total number of establishments becomes smaller. In a period of very rapid growth, it may happen not only that each unit becomes larger but that the total number increases. More commonly, however, the total number decreases or at best remains stationary while yet the individual establishment becomes greater in size, and the productiveness of the industry as a whole is much enlarged.

The causes of the growth of large-scale production are to be found mainly in the revolutionary changes in the arts during the last century and a half. Underlying them all is the increasing division of labor and the increasing use of machinery. A necessary condition—tho not in itself a cause—has been the widening of the market under the influence of cheaper transportation.

A tool or machine of any kind is advantageous only if used repeatedly for the same operation. The greater the number of repetitions, the more is it worth while to have an elaborate tool and to give much labor to its making. Machinery moved by power is a highly elaborate tool. The larger the scale on which an enterprise

is conducted, the better is the opportunity for using machinery to advantage. The gain from its use arises from several sources. Power itself becomes cheaper per unit as it is applied on a large scale. Both for first installment and for running expenses a large steam engine costs less, for each horse power, than a small one; which means economy if the establishment is large enough to utilize all the power supplied. Again, subsidiary operations can be carried on to advantage by machinery. The use of steam shovels in handling coal, ores, earth, and of similar instruments for loading and unloading vessels, depends on the work being massed in large volume at one spot. An ocean steamship of 10,000 tons carries freight more cheaply than one of 5000, and one of 20,000 tons more cheaply still. Wherever the traffic is heavy, as between Europe and the United States, the huge steamship is economical. Where the traffic is less heavy and less regular, the ship of moderate size holds its own. The largest of the American corporations making agricultural implements, one that illustrates conspicuously the tendency to large-scale production—the International Harvester Company—has a machine whose sole work is to shape poles for wagons and harvesters. The machine cost $2500; it saves a cent per pole; it is worth while only because poles by the hundred thousand are made each year.

Other causes, more or less closely connected with the growing use of machinery, have strengthened the tendency to large-scale production. Just as all the several expenses for the plant and power become less per unit as the output enlarges, so the general expenses for administration and counting-room work tend to become less. Clerks are kept more continuously occupied and more elaborate division of labor among them is feasible. Superintendents and foremen can take charge of the full number of men which each can direct to advantage. One watchman, one engineer, one timekeeper, can usually serve a large establishment as effectively as a small one. All the miscellaneous expenses of general management are less in proportion to a large output.

The mercantile management of a large enterprise—the buying of materials and the selling of the product—also offers opportunity for economy and efficiency. Supplies can usually be bought to

greater advantage. This is commonly spoken of as if simply the result of greater bargaining power on the part of the large buyer and greater pressure of competition among those who wish to sell to him. But in the main it arises from the fact that mercantile operations themselves, and especially wholesale operations, are carried on more economically when on a large scale. Expenses for clerical work, rentals of office premises, and the like, which constitute the main outlays of the wholesale dealer, are no greater for large transactions than for small. Hence brokers and wholesale dealers can sell at lower prices to those who buy habitually in large amounts.

Again, the disposal of the output is often less expensive per unit for a large establishment than for a small one, and often at still less expense for a very large establishment than for a moderately large one. Advertising and notoriety much affect the marketing of sundry commodities. When once appeal is made not to a limited local market but to a large and extensive constituency, the disposal of the great quantities of goods turned out by a modern factory becomes by no means the least difficult of its manager's tasks. All the apparatus for drumming up custom—traveling salesmen, trade catalogs, and the like—is the more effective, and the less costly per unit of product, in proportion as it operates on a large scale. Advertising is most effective when spread over the land with every sort of device, when it is systematized and put in charge of a separate manager. All such elaboration of marketing is both a result and a further cause of a great volume of business.

The utilization of "by-products" [1] is another of the advantages of large-scale production. At the great packing houses which do so much of the butcher's work of the United States, every particle of the slaughtered animal is used, and many things which would go to waste in the small shop become a source of profit. A very large woolen factory finds it advantageous to utilize the fatty matter which is attached to the wool as it comes from the sheep's back. This grease, which must in any case be scoured out of the wool, goes to waste in a smaller establishment; whereas the large mill, by putting in a plant for the special purpose of treating the

[1] Better, "joint products"; see Chapter 15, § 1.

grease, finds it a source of gain. Great ironworks find it possible to utilize the gas expelled from coal in the coking process; either selling the gas, purified, in a near-by city or using it at once for fuel in their own furnaces. A large sawmill can put in a plant for burning its own sawdust, dispensing with other fuel for power.

Other advantages of large-scale production arise from the possibilities of experimenting with new devices and new methods. Some ventures will fail, some succeed. In a very great enterprise, the successes may be expected in the long run to outweigh the failures; the enterprise insures itself, so to speak, against the inevitable risks of experimenting. Where the operations are conducted on a small scale, the failure of one experiment may ruin the entire undertaking. Again, the best technical skill, the best-trained engineers and chemists, are more easily and more economically employed by the great establishment. As with expensive but efficient machinery, their use is advantageous only for a very large output, and is most economical for the largest output.

A curious case, instructive both as regards the possibilities of economy from large-scale operation and as regards the drift of modern times, is that of retail dealings. Retail trade long seemed to be the special field for businesses of small or moderate size. But large-scale retailing began in the nineteenth century in the larger cities and became more and more effective with the improvements in urban passenger transportation, a striking case being that of the so-called department stores. A farther movement in the same direction came with the development, even more striking, of the chain stores and the mail-order houses. The small retail shop of the older days had the advantage—or at least had one great advantage—in that the purchaser's source of supply was near. But the chain store is as handy as the small retailer, and has reached nooks and corners which before were inaccessible except to him. The motor truck was no small factor, since it enabled the scattered stores of the chain to be continuously and cheaply supplied with their stocks from central stations. Next came the mail-order concern, made possible in its turn by the facilities which the post-office was able to provide for the shipment of parcels. Both the chain store and the mail-order concern needed a technique of their own. New

systems had to be devised for standardizing the articles, keeping the accounts, supervising the staff. The growth in both directions was in the United States extraordinary, and probably far surpassed the dreams of the first experimenters. And in both the dominant element was large-scale management rather than large-scale production in the narrower sense—a distinction about which more will be said presently. In the opening years of the twentieth century the small dealer seemed to be still holding his own tenaciously in the merchandizing of goods. But by the time the century had run one third of its course, the large concern had made great conquests here also; its triumphs resting not only on mechanical invention but the resourcefulness and incessant scheming of the business manager.

§ 2. The limitations on large-scale production arise mainly from the infirmities of human nature. The extension of the scale of operations means an ever-increasing reliance upon hired labor an an ever-lessening reliance on spontaneous self-interest. If all men worked with as much energy and spirit for an employer as they do for themselves, the spread of large-scale production would be almost without bounds. A striking illustration of the influence of this limiting factor is shown in the differing tendencies of agriculture and of manufactures.

The operations of agriculture are necessarily spread over a considerable area and they are not easily subjected to a fixed routine. Both circumstances make supervision difficult. Manufactures, on the other hand, bring the concentration of hundreds or thousands of workmen under a single roof or in a small area. Moreover, in manufactures, machinery means the repetition of identical operations. Hence a routine can be fixed, and workmen assigned to fixed tasks, and their faithfulness controlled, with comparative ease. But in agriculture much must be left to the zeal and intelligence of the individual worker.

The consequence is that agriculture has nowhere shown the same tendency to enlargement of the scale of production which is so unmistakable in manufactures. It is true that some countries are usually spoken of as countries of large farming; England is the type of such a country. It is true, also, that in some parts of the

United States (in the North Central region, for example), there has been in recent years a slight tendency to increase in the size of farms. But a farm which is called large is an industrial unit of comparatively small size. One which employs twenty men the year round is considered large, yet a factory employing this number is a small affair. The tasks of twenty men engaged in farming would be spread over several hundred acres, and must present troublesome questions in assigning and supervising the work. Farms of this size are comparatively rare. By far the greater part of agricultural work is done on farms where a single man, having under him perhaps one other or a few others, conducts the operations on his own account. In the early stages of the development of some parts of the United States, so-called "bonanza" farming appeared for a time. Where great level tracts of fertile land were suddenly opened to cultivation, as in the interior valleys of California or in the Red River Valley of the Dakotas, wheat culture was carried on for a while over thousands of acres, with dozens, even hundreds, of men and vehicles and with expensive machinery. But this proved only a temporary phase. As the fertility of virgin soil began to be exhausted, and a more varied and careful use of it was called for,[1] these great tracts were split up into smaller units. The head of a large factory can devise means for supervising his men and for securing the execution of his orders. But the owner of a farm can use hired labor to advantage only when his own example and his own oversight supply the needed stimulus.

Some industries, tho spread over a large area and presenting difficulties for the supervision of hired labor, are so much more effective when on a large scale that these disadvantages are not decisive. The railway is an example. Many of its employees are necessarily scattered over great tracts of country. The supervision of the innumerable agents calls for an intricate and expensive apparatus of rules and regulations, bookkeeping and auditing. But the work is done so much more cheaply on a large scale that this difficulty and the expense entailed by it are more than offset.

A glance at such a volume as the *Statistical Abstract of the*

[1] Compare Vol. II, Chapter 44.

United States, with its summary of the number of establishments and volumes of transaction in various kinds of business, shows instructively which among them, for reasons of this sort, resist the tendency of concentration. The strictly manufacturing establishments show the characteristic features of the modern movement. Tho the volume of transactions becomes immensely greater, the number of establishments becomes less. So it is with the manufacture of agricultural implements, of boots and shoes, of carpets, chemicals, firearms, glass, cotton, woolen and silk fabrics, sewing machines. In those industries which purvey more directly to the consumer, or for other reasons must be near the persons with whom they have dealings, the number of establishments increases in proportion to the increase in population and the volume of their own transactions. Such are blacksmithing, carpentering, plumbing, printing, painting, and paper hanging. Here there is no marked tendency toward an enlargement of the size of the individual establishment, still less any complete victory of great-scale production.

§ 3. The limitations of men's faculties explain why large-scale operations do not make their way, even in manufactures, with unfailing certainty. What has been said in the preceding paragraphs may seem to imply that the transition to greater size takes place automatically. This is by no means the case. It depends on the energy, ambition, insight, of individual men. Every new machine and every change to larger scale involves risks, calls for planning and judgment, is dependent on some individual's initiative. If an indefinite number of individuals were capable of this sort of work, the march of progress would be faster and large-scale operations would make their way more surely and speedily. As it is, these changes wait on the impulse given by the comparatively few who have the capacity for industrial leadership. Time and again some such individual reorganizes his business upon a larger scale and with more highly developed plant and machinery. Then others follow his lead, and a whole industry is rapidly transformed. This has happened during the last two decades in the iron manufacture, especially in the United States and in Germany. Carnegie in the former, Krupp in the latter, led the way in a remarkable

development. Usually, however, the advance takes place by more gradual and tentative steps, like those in the growth of the size of ocean steamships. The industrial revolution, as regards its pace, has been in reality not a revolution but a slow and gradual change, dependent on the energy and ingenuity of individuals and limited by the scarcity of men possessing such qualities.

This human factor is usually ignored by the socialists and the constructors of utopias. To them it appears that the increase of productive capacity is a simple matter, and simplest of all in manufacturing industries. Double or quadruple the scale of the individual establishment; shut up the small ones and transfer their workmen to those of large size—here is a ready way of increasing output and making matters easy for all mankind. It is part of the same utopian attitude that unending improvements in mechanical appliances are supposed to be an assured asset of the future under any condition of society. The fact is that great advances in the arts, whether they involve fresh invention or better management, or (as is commonly the case) involve both of these factors, arise from individual initiative, individual calculation, individual leadership. In a socialistic state, as under a régime of property, a fundamental question will be how men shall be induced to scheme and invent, to improve and to perfect their faculties to the utmost. What motives now move men thus to open the paths of industrial advance, what other motives may conceivably actuate them under different social conditions—all this must be reserved for later consideration. Under no organization of society is there any royal road to boundless increase of production.

§ 4. A new phase of large-scale production, already alluded to, has come to be of great and almost ominous importance during the present generation. Perhaps it should be called large-scale management rather than large-scale production; since it involves not so much an increase in the size of the individual establishments as the combination under single management of several establishments not greatly different in size. It takes two forms, which may be described as horizontal and vertical.

Horizontal combination is the union under single management of a number of enterprises of the same sort. They are usually few,

and each is usually on a large scale. There is, to be sure, the striking case of the chains of retail stores, where the individual establishment had shown no marked tendency to enlargement under separate management, and still shows none under combination and large scale management. But the more common case is that where the size of the representative establishment in an industry enlarges, the number of individual establishments shrinks, and the stage is finally reached where but a few survive—a dozen, perhaps. These then combine; not in the sense that one huge establishment supersedes the dozen, but that the dozen, while retaining their technical independence, are owned and managed as one. Tho large-scale operation may have reached its limit so far as the mechanical apparatus of production goes, some gain may still be secured from united large-scale administration. A typical example is the American Sugar Refining Company. A modern refinery is a huge concern, costing a couple of millions of dollars, and putting out 10,000, even 15,000, barrels of sugar a day. Yet there are limits to its size. Beyond a certain point, enlargement no longer adds economy in operation. When an output beyond this capacity is called for, a second refinery of the same kind is erected, and so on until the total supply is provided. All these refineries, however, may be managed from one common center, with at least possibilities of resulting economy. Their supplies may be bought in common, and distributed among them in such a manner as to insure continuity in operation and the minimum outlay for transportation. This last factor, economy in transportation, is of great consequence, where the chief material (raw sugar, in this instance) comes from great distances and, being rapidly worked up, must be continually and systematically replaced. Machinery may be made identical, or "standardized," in the different works, and its repair and replacement thus facilitated. These and other possible economies may be offset, to be sure, in whole or in part by the inherent difficulties of large-scale management—notably the increasing difficulty of supervision. Experience and especially the test of competition can alone settle with certainty whether the advantages offset the disadvantages.

Horizontal combination is typical of the so-called "trust." The

motive for such union under single management is twofold. Partly it is to secure economy in management, but largely it is to put an end to competition and bring about a more or less effective monopoly. So far as economy is secured, the movement may be to the public advantage. But if monopoly develops, it has public disadvantage. How far monopoly in fact is likely to result, and how far cheapening of production is in fact brought about, is still uncertain; time and experience alone can show. But it is clear that in some respects at least, and for some industries, such combination brings an extension of large-scale operation and concentrated management.

Different in its essential features is vertical combination or, as it is sometimes called, the integration of industry. The usual outcome of the division of labor has been that the several steps in production which succeed each other in time have been conducted in independent establishments. But in some important trades there has appeared a tendency to unite such successive stages under single management. Thus the iron industry, in the older form of organization, was split up into a number of separate branches. One producer—that is, a capitalist hiring and directing a group of workmen—carried on ore mining, and disposed of his ore to other producers engaged in smelting it into pig iron. Still another producer similarly cut the wood and converted it into charcoal—this in earlier days, when wood supplied the fuel for iron making—or, after coke supplanted charcoal, mined the coal and made it into coke. The pig-iron maker, who had bought the ore and the fuel, sold his product to the puddler or steel maker, who in turn sold his bar iron or steel to the machinist, the builder, the wire maker. Vertical combination, or the integration of industry, appears when all these successive steps are united under single management—when many of these phases of iron and steel making are combined in one great enterprise.

The United States Steel Corporation carries out this sort of combination in a typical manner, and on an enormous scale. Itself a union of previous combinations which had adopted the same method on a scale already great, this corporation owns vast mines of iron ore, of coal, and of limestone. The iron mines are situated

chiefly on the shores of Lake Superior, the coal mines chiefly in Pennsylvania. A large part of the ore is carried to the coal, and smelted in the great iron-making district of which Pittsburgh is the center; but in part the coal is carried north and west, meeting the ore, to be smelted at various places on the Great Lakes. To transport these materials, the corporation has its own railways in the Lake Superior region and in the region from Pittsburgh to Lake Erie; and it owns a great fleet of steamers and barges on the lakes. The pig iron, made in its own furnaces, is converted into steel of various shapes in its own steel mills. The further operations of converting the steel into rails, structural and bridge shapes, plates and sheets, tubing, and wire are carried on in still other establishments. In no other industry, and nowhere else in the world, has vertical combination been conducted on so great a scale.

The iron and steel manufacture offers an unusually tempting field for such combination, chiefly, it would seem, because of the concentration of the supplies of raw material—coal and iron ore. Those who, at any stage of rising demand, possess the mines of coal and iron, have the whip hand in the situation; hence the manufacturers of the more finished forms of iron and steel have sought to gain control of the mines, by purchase or amalgamation. This tendency has shown itself in some degree in Great Britain, and has proceeded in Germany almost as far as in the United States. A series of superimposed establishments has now become the normal form of organization in the iron manufacture.

Some tendencies of the same sort are found in other industries. The International Paper Company owns great tracts of spruce forest, cuts the timber and logs, floats them to its own pulp mills, and there manufactures the paper which is used in such enormous quantity by our newspapers. The Harvester Company, already referred to, owns forests and cuts timber; it owns its iron and coal mines, and makes its iron and steel. The Sugar Refining Company owns its forests and makes its barrels. The Diamond Match Company owns and manages large forests for assuring a supply of the wood best suited for the billions of matches which it annually turns out. The Goodyear Company operates rubber plantations in

the tropics, and also cotton plantations in the United States, as well as textile mills in New England and elsewhere, to produce the special cotton fabric needed for the modern rubber tire.

Other industries have shown a similar development in another direction—in the marketing of goods. The usual arrangement is for a separation between manufacturing and marketing. The shoe manufacturer commonly sells his output to the wholesale dealer or "selling agent," who in turn often sells to an intermediate dealer, the jobber, and sometimes directly to the retailer. But some shoe manufacturers have undertaken not only the making but the marketing of their wares. They have established their own retail shops, scattered in many cities over the country, and thru these deal directly with the consumer. Again, the American Tobacco Company, by establishing its own retail shops in great numbers, likewise combined the distribution of goods with their production.

Vertical combination and horizontal combination may go hand in hand. The American Tobacco Company attempted to combine all the establishments manufacturing tobacco for smoking and chewing; and the extension of its operations into the retail disposal of its products has been the outgrowth of the endeavor to form and strengthen this all-embracing horizontal combination. The Steel Corporation owns many iron furnaces, many steel mills, many tube works, many sheet-steel and tin-plate works, and thus exemplifies also the union of the two kinds of combination. In Germany, the *Stahlwerksverband* (Steel Works Association) has formed a compact pool in the iron and steel manufacture, tho one that does not go the full length of completely unified ownership. In Great Britain, on the other hand, while many large works have extended their operations downward to the control of mines and upward to the making of finished products, there is much less of horizontal combination; the several great enterprises go their own way independently.

The movement toward vertical combination is less widespread than that toward horizontal combination. The iron trade, which presents so striking a case of the former, is after all an exception. The desire to secure control of a limited or at least concentrated

raw material, which has promoted the integration of the iron trade, has not affected others in which the sources of raw material are more scattered. In the manufacture of cotton, wool, silk, or flax there is no indication of any movement for control of the supply of raw material or for widespread vertical combination in any other way. On the contrary, the tendency seems rather toward a minuter division. The textile industries in Great Britain and on the Continent have always been split up into separate industries to a greater degree than in the United States. In Europe spinning, weaving, bleaching, dyeing, printing, are usually carried on as distinct industries. The tradition in the United States for a long time was for the combination of several of these steps—especially spinning and weaving—into one organization; then, in the early years of the twentieth century, there was a movement in the other direction. In the shoe manufacture, while there has been the marketing arrangement just noted, and in some cases a combination of leather tanning with manufacturing, the trend does not seem to be toward great combinations of the vertical type, whether reaching backward to the supplies of material or forward to the marketing of shoes. The experimental and tentative character of this phase of industrial development is indicated by such differences in trend—one way in the United States, the other way in Great Britain. It is all a matter of trial and error, judgment and venturing; there is no certain balance of gain, nothing fatalistic.

Combination, whether horizontal or vertical, is in part a result of the intensified competition which comes with the greater investment of fixed capital and the greater size of the separate enterprises. But very largely it results from the discovery of the possibilities of organization. What are the limits to the size of the enterprise which can be managed as a unit? The single factory or shop, perhaps large, was supposed until comparatively recent times to represent that limit. But as the scale of industry has been enlarged, the operations have been systematized and subjected to more perfect control. The task of management itself has been subdivided. Separate persons are intrusted with the purchase of supplies, the sale of product, the maintenance of plant, the hiring and superintendence of labor, accounting and auditing. The

genius of men with great inborn capacity for business has led to ever-greater perfection of organization. Technological changes in ancillary industries have played an important part; the telegraph, the telephone, improved postal service have promoted large-scale management as they have large-scale production. Yet on the whole these striking changes have resulted mainly from the skill, judgment, and administrative capacity of the guiding individuals.

None the less, the larger the scale of operations the more do its disadvantages appear. There is need for an expensive system of control—for supervision, accounting, auditing, the effective prompting of energy and economy. The test of competition settles in the long run whether the great combination is the more efficient agent in production. If it can produce more cheaply, it can sell more cheaply and displace its rivals.[1]

§ 5. Notwithstanding the wastes of competition, and the possible economies of large-scale production, competing establishments hold their own over the greater part of the field of industry. There is no present prospect that competition will be everywhere supplanted by combination and monopoly.

That competition operates wastefully seems in some cases obvious. The milk of a city, for example, is usually supplied by a number of dealers, each with his own set of customers scattered irregularly over a large area. If all who lived in a given quarter were supplied by one dealer, a clear economy in delivery would be secured. If the whole supply for an entire urban district were under single large-scale management, there would be a possibility of cheapening the product still further and (what in this case is specially important) of improving its quality. In this case competition has so many disadvantages, and regulation is called for on so many grounds, that monopoly would seem desirable as well as inevitable. In such things as groceries and foodstuffs, also, retail dealers overlap in similar wasteful fashion. Commonly, too, the areas supplied by competing manufacturers overlap. Advertising is in large part designed to induce a customer to turn merely from one dealer to another. If there were no competition—if one large

[1] To this statement of the automatic action of competition there are qualifications, considered in Vol. II, Chapter 65, § 3.

establishment supplanted ten small rivals—the same wants would presumably make themselves felt, the same purchases would be made, the expense of advertising eliminated, the goods produced and marketed more cheaply.

Tho some tendency is seen toward getting rid of the causes of waste, the tendency is not very marked. With the growth of great cities, large firms and companies have come in great degree to control urban milk supply, yet with little indication that complete and systematic combination is emerging. In other industries, the great manufacturing "trusts" endeavor to avoid cross freights by making shipments from that one among their establishments which is nearest the point of delivery. But as a rule manufacturers continue to compete and to ship in a seemingly haphazard way. The same is true of retail trade, where all sorts of establishments, great and small, vie for the customer and duplicate facilities in the traditional and apparently wasteful fashion.[1]

The waste, however, is less than it seems. Competition keeps every one keyed to a high pitch, nerves the shrewd and alert, weeds out the inefficient. Advertising is part of the mechanism of competition as well as of combination. Not least, competition leaves the purchaser some freedom; he is not subjected to the alternative of either turning to one single purveyor or else doing without. Even the most benevolent and considerate monopolist often becomes exasperating; how much more so the ordinary trader when no longer spurred by competition! A choice as to what you would have, and when and how you would have it, satisfies a deep-rooted human desire. But the consumer in the socialist state would have to accept whatever the all-controlling public managers put before him. The satisfaction which comes from freedom of choice explains in large part the persistence of competition.

[1] Compare, however, what is said below on monopolistic competition, Chapter 18.

CHAPTER 5

CAPITAL

§ 1. Production is spread over time. This fact disguised by the division of labor. Increasing use of plant and machinery in modern times.—§ 2. Producer's wealth and consumer's wealth; capital.—§ 3. Capital rests on a surplus.—§ 4. In what sense capital rests on saving. Hoarding contrasted with saving for investment.—§ 5. Investment means advances to laborers. Inequality of possessions in relation to advances. Middlemen for investment and advances.—§ 6. The maintenance of capital, as well as its creation, involves saving.—§ 7. Some difficulties of terminology.

§ 1. The increasing complexity of the division of labor and the growing use of machinery have added to the number of separate stages in production and to the length of time over which the whole process is spread. Hence the greater need of a supply of tools and materials, the importance of capital, the problems which relate to owners of capital and to the income from capital.

Production is spread over time in any society advanced beyond the most primitive savagery; and this not merely for the several subdivided steps in production but for production as a whole. That agriculture takes time, from the sowing of the seed to the reaping of the crop, is obvious. But the sowing is not the beginning, nor is the reaping the end. The seed must have been itself sown and husbanded, and the tools for cultivation must have been prepared in advance. After the harvest the grain which is reaped may indeed be available for satisfying human needs almost at once; it is so in a small, self-contained community, such as we still see in a village of Hindustan. But in the countries of advanced civilization grain is carried by rail or water to a mill, probably distant; there is ground into flour; then carried another distance to dealers; and finally, after a considerable interval, put into the hands of the consumer. Each of these steps not only takes time in itself but requires apparatus which has been made in the past and has taken time to make—the railway or steamship, the flour mill, the warehouses and shops of the middlemen. Almost all the operations of production involve first the procuring of materials from

nature's resources, then their fashioning with the aid of tools and machinery. Let the reader but consider the mode in which the familiar articles of daily use have come into his hands—the clothing and the footgear, the furniture and household utensils, the books and ornaments, the house in which he dwells—and he will see how long has been the series of operations, how intricate the division of labor for each one, and how extended the period from the beginning of production to the final attainment of the consumable or enjoyable article.

This fundamental fact, resting on the complex division of labor, is yet disguised by that very division. The tanner who puts his leather on the market, the farmer who sells his flax, the ironmaster who sells his steel or iron, each thinks of himself as marketing a completed product. By the sale he gets money, and so the command of the enjoyable things he wishes to buy or of the things needed for continuing production. He never stops to reflect what must further be done to the thing which he sells; how it must pass thru the hands of a long chain of producers and dealers before it reaches in consumable form those whose wants are finally satisfied.

In modern times, the most significant aspect of this element of time in production is found in the increasing use of machinery and plant of all sorts. Machinery, tho it may be simply a more intricate kind of tool, adds so much to the preparatory work that it has greatly accentuated the problems that arise from the spreading of production over time. A factory requires a year or years to build; the machinery in it requires still more time to make. Many years are needed for constructing a railway; a generation for such a work as the Suez Canal or the Panama Canal. The factory and the machinery in it exist for the purpose of eventually turning out things to be used and enjoyed. The railway and canal facilitate the geographical division of labor and serve to promote, thru a series of further steps which begin only when these means of transportation have been completed, the eventual abundance of things to be used and enjoyed. One simple fact illustrates how marked the tendency toward greater use of plant has been in the period since the industrial revolution began. The world's annual production of

iron has multiplied tenfold the last half century, and hundredfold in the last century.[1] Iron is used solely (the exceptions are insignificant) as an instrument of production; it is the foundation of the material apparatus of civilization; it means plant, tools, machinery. The enormous quantities of it which have been turned out in modern times signify an extraordinary increase in the construction of elaborate and expensive apparatus, and a corresponding extension of time in the operations of production.

§ 2. If we were to take a cross section of the community's possessions at any given time, we should find them to be of the most diverse kind. There would be, in the first place, such things as iron ore and steel bars, timber and wool and cotton, factories and railways and ships, stocks of all sorts in warehouses, commodities ready for sale in the retailers' shops. And in the second place, there would be houses, furniture, clothing and food, in the hands of those using them for the satisfaction of wants. To the first set of things we apply the term capital, or producer's capital; the second set we call consumer's goods, or wealth that is not capital. The first set we may speak of as unfinished goods, the second set as finished and enjoyable goods. For some purposes of economic analysis they are similar, for other purposes dissimilar. The difference between them is essentially one of degree, yet is so great as to justify a distinction.[2] For the present, we shall find it convenient to apply the term "capital" specifically to the first set—to producer's capital. The second set will be referred to as enjoyable or consumable or finished commodities; and only when speaking of them in those aspects and relations which offer analogies to the first, shall we refer to them as consumer's capital.

[1] The world's annual output of pig iron was:—

In 1800	825,000 tons
In 1850	4,750,000 tons
In 1870	11,900,000 tons
In 1910	60,500,000 tons
In 1930	79,000,000 tons
In 1936	91,000,000 tons

[2] The difference in degree is one as to the time when satisfaction or utility accrues. That time is commonly nearer in the case of consumer's goods, and more distant in the case of producer's capital. Compare what is said in Vol. II, Chapters 38 and 47.

Capital, then—producer's capital—is not in enjoyable form; it is not now a source of satisfaction. It exists for the purpose of increasing consumer's wealth. Its relation to enjoyable goods is twofold. On the one hand, it may be said gradually to "ripen" into such goods. On the other hand, it is a means of increasing their supply.

It is easy to see that raw materials, as they are commonly called, ripen into finished commodities. Wool is converted by successive steps into clothing, grain into bread, stone and timber into a house. A process the same in essentials takes place with tools and machinery. Suppose a printing machine to last for one year only, being worn out and worthless at the close of the year. The books printed with its aid are the product not only of the labor applied to making the paper and other materials, and of that of the compositors and other workmen in the printing office, but also of that involved in the construction of the printing machine itself. If we suppose that one hundred books are printed in the course of the year, the machine may be said to have ripened into so many enjoyable goods, and each of these may be said to have embodied in it one hundredth of the labor which was given to constructing the machine. The machine as such disappeared, just as the paper and ink as such have disappeared; in place of all three we have the printed books. If the machine lasts for ten or twenty years, the labor of constructing it contributes to making a much greater quantity of books, and a smaller fraction of the labor of construction is embodied in each book. So of all machinery and all plant. It wears out sooner or later, and may be said sooner or later to ripen into goods that satisfy our wants.

The most important single cause of the abundance of consumable goods, and so of the improvement in the material welfare of mankind, is found in those forms of capital which are commonly spoken of as fixed—in tools, machinery, plant. Certainly this has been the most important cause of the remarkable advance in material welfare which the civilized countries have made during the last two centuries. Erect a great cotton or woolen mill, a shoe factory, a large sugar refinery or flour mill—take much time and apply much labor for getting ready an elaborate apparatus—and

eventually you will secure your product in greater abundance, and with less labor embodied in each unit. The making of machinery itself has illustrated this tendency as strikingly as any other branch of production. The manufacture of iron and steel, conducted on a great scale with elaborate and expensive plant, serves to turn out in cheapness and abundance the metal indispensable for all the varied instruments of production. Locomotives, textile machinery, agricultural implements, not to mention the simpler tools of the mechanic, are themselves made with machinery.

In order that all this application of plant may work smoothly and effectively, the supply of materials must also have been on a large scale; and this again involves prolonged preparation. A great iron furnace, kept in blast night and day, year in and year out, takes into its maw huge quantities of iron ore, coal, and limestone. These, no less than the furnace itself, must be made ready in advance. So the textile mill requires its wool or cotton or silk, the shoe factory its leather, the refinery its raw sugar. Thru all the complicated operations the trend is the same; elaborate preparation, production spread over time, much capital, eventual plenty and cheapness of the consumable goods.

§ 3. In order that there shall be capital and time-using production there must have been at some previous period a *surplus*. The more of capital there is to be, the more must there be a surplus to draw on.

In the very earliest stages of the formation of capital, that surplus showed itself in the fact of spare time. The first rude implements of stone and bronze must have been fashioned during hours when labor did not need to be given for the satisfaction of imperative wants—when there was a chance of doing something else. What motives may have influenced man during this stage, and by what chance the first tools were hit on, we cannot guess. Possibly a mere instinct of contrivance was the moving force. A reasoned understanding of the gain from having tools and supplies must have set in at an early stage. The choice under the simplest conditions is between the present and the future—between idleness or amusement for the moment and provision for future needs.

The greater the surplus, the greater the time and labor which

can be given for future needs. When the arts are at so low a stage that little is produced beyond the bare necessities of existence, provision for the future can be made only on a scanty scale. On the other hand, the very scantiness of capital is an obstacle to the effectiveness of labor and so to the existence of any considerable surplus. During long ages mankind was thus in a position of double difficulty. Without capital the productiveness of labor was meager, and yet with meager productiveness of labor there was little possibility of creating more capital.

It is not to be understood that the slenderness of the surplus stock was the only obstacle to the creation of capital. Ignorance of natural laws and of the possibilities of tool making, indifference to the future, were no less important. But without the surplus the very foundation for building up any effective apparatus of production was lacking. Here, as often, the first step was the hardest. Once man had become possessed of some capital the productiveness of his labor became greater and thereby the creation of still more capital became easier.

§ 4. In the preceding section capital has been spoken of as made or created. But capital is also said to be *saved* and accumulated. Both expressions are permissible. If we think of one person or set of persons as being alone concerned with the several steps by which capital comes into existence, we can see that this person not only provides for the future by saving but also uses his surplus in shaping tools or getting together materials. In a society having an elaborated division of labor, the two things are rarely done by one person; that is, they are rarely done together by one person for any given item of capital. A machinist may save, but there is no connection between his present savings and his present work on the machines. It is the previous savings of other people that made possible the materials and the machines as he finds them. When all incomes and expenditures take the form of money, savings are made not by putting aside things in kind for one's own use but by putting aside money for future needs. On the other hand, tools and other apparatus of production are made for the market by persons who are not consciously providing for the future. They are then bought by other persons who wish to

"invest." The process by which these separate steps are made to bring about their joint result in the modern organization of industry deserves careful consideration at this stage; more will be said about it as we reach the topic of interest on capital.

Saving may take the form of simple hoarding. The miser who puts away a store of coin, saves and provides for his own or other's needs. But no addition to the apparatus of production results from such saving. Where property is insecure, from the rapacity of a despot or from the feebleness of a government unable to protect against foreign invaders, hoarding is sometimes done on a large scale. In British India, during many centuries preceding the British occupation, both these causes of insecurity existed. Hence persons who had means put them largely into the form of specie and jewels, articles having much value in little bulk and capable of being hidden or carried away. The European aggressors of the seventeenth or eighteenth century found great stores of such wealth in Hindustan, not because that country had rich mines but because the people had attained a considerable civilization and prosperity and had hoarded long. Notwithstanding the peace and security which British rule has long maintained, the habit of putting accumulated means into this form has continued in India to our own time. In France, for a long period preceding the French Revolution, the peasantry—those among them, comparatively few, who had anything at all in the way of a surplus—put away coins one at a time, hidden in the chimney or garret until enough had been accumulated to buy a scrap of land. Fear of spoliation and ignorance of other ways of doing anything with the money caused their saving to take the form of hoarding. No addition to capital was thereby promoted. Nor was there any addition to capital even when the accumulated coins were brought out for the purchase of land. The noble from whom the purchase was made probably frittered away the proceeds, and the only immediate result of the peasant's accumulation was the transfer of land from one hand to another. Such practices continued in France after the Revolution and indeed thru the nineteenth century. The Franco-German war of 1870–71 and still more the World War of 1914–18, leading as they did to enormous and widely diffused

public borrowings, finally made a great breach in the peasants' habit of putting aside hoards of specie.

The great bulk of saving, however, takes in modern times the form of *investment*. Contrast the process of hoarding with what happens when money is put away in a savings bank—an operation which we may select as typical of the process of investment in a modern community. The person who leaves his cash with the savings banks commonly thinks only that it is safe, and that he is paid something as interest on it. But the cash is not kept in the coffers of the institution. A small fraction only is retained, to meet possible calls of depositors who wish to make withdrawals. Almost all of it is lent out to persons who use it for making a profit. Now profit arises in the ordinary course of things from the operations of production; and the person who borrows money uses it for the purchase of things needed in production. He may be a manufacturer who erects a building, buys machinery and supplies, hires workmen. He may be a merchant who buys commodities from the manufacturer and carries them one stage further in the successive stages which bring them at last to the consumer. Every person who directs production—such as the manufacturer or merchant—uses a large part of his means in buying materials or tools or stores from producers of a previous stage, so recouping these for the outlays they have already made. The money means which are put at the disposal of the business class as a whole are the most important part of the mechanism for adding to the concrete apparatus of production.[1]

§ 5. The fundamental fact in this elaborate mechanism of saving and investment is that advances are made to laborers. One set of persons puts aside money means; thru various channels other persons are given command of these money means and use them to set laborers to work. Here, again, the division of labor between those who carry on the successive stages of production conceals the essential nature of their operations. A manufacturer spends only a part of his means upon hiring laborers directly; the

[1] This must be taken as a preliminary statement. There is much more to be said about the relation between saving and investment in the operation of banks other than savings banks, and indeed about the whole capitalistic process.

rest he uses in buying plant and materials and in the other expenses of production. But those materials were themselves fashioned by laborers to whom another set of advances had to be made by a previous capitalist. The wholesale or retail merchant hires comparatively few laborers—only a set of clerks and a porter or two. But by his purchases of goods he recoups the advances of a long series of preceding employers, and he himself is concerned only with the finishing touches in the whole process. Looking at the operations of capitalists and employers as a whole, and analyzing the outcome of the division of labor among them and their workmen, we find that all capital is made by labor, and all the operations of the capitalist class are resolvable into a succession of advances to laborers.

These advances, just spoken of as money turned over to laborers, consist ultimately in a provision of commodities for their use. The money is but the medium whereby laborers get command of the commodities which they buy. These commodities—things to eat, to wear, to give shelter—are in the last analysis what the employing class hands over to those whom it employs. Some of the advances were made in the past and are represented now by plant and materials still in use, of which the full equivalent has not yet been reproduced in finished form. Some are made from day to day, in the course of current operations. The whole of existing capital may thus be described as a great accumulated surplus which has been used and is being used for maintaining labor while provision is made for the future. The process of setting laborers to work in the initial stages of production is going on all the time; similarly the process of bringing articles to the final stage of consumable form.

The wide separation, in modern societies, of the two acts needful for the creation of capital—saving and the application of labor—is mainly the result of inequality. Persons of the well-to-do class have a considerable surplus over current needs and save with comparative ease. They own most of the apparatus of production. But in our modern societies the great majority of men are not of the well-to-do class and have little in the way of a surplus. They have small accumulations, and are mainly hired by others in carrying

on the operations of time-consuming production in making and maintaining capital. No doubt some savings are made by the working classes, and thru the agency of savings banks and similar institutions these savings have increased rapidly. But while absolutely considerable, they are no large proportion of the total accumulated means. The greater part of the capital owned and maintained in modern communities arises from the savings of the comparatively small number of the more fortunate classes.

A chain of middlemen commonly connects the individual who saves with the laborer to whom advances are made. The employer himself, tho he almost always uses some means of his own, usually is a borrower. He borrows, however, not from the savers directly but from their various agents and representatives. The savings bank, for example, collects surplus sums from individual savers, yet often deals with the employer of labor only thru brokers and other middlemen. It buys stocks and bonds from the broking and banking firms. Such firms issue them after long negotiations with the persons undertaking the operations to which the whole series of transactions is in the end directed. Bankers are the typical intermediaries; their essential function is to direct the stream of available money income into one direction or another, and to put into the control of employers the means for setting laborers to work. Life insurance companies, which collect funds put aside by many individuals in order to provide for future needs, are among the great modern agencies of saving. Like the savings banks, they commonly make their investments not by direct loans to employers but thru bankers and other intermediaries who take the first risks of production and aim to bring the investors a secure return. In modern times there has been a great development of the class of middlemen who intervene between the investors and the active managers. There have been great possibilities of profit for those middlemen, great possibilities of abuse in positions of trust, but also great effectiveness in collecting and investing the savings that underlie the enormous growth in the total capital of modern communities.

§ 6. Not only the creation of capital involves labor and saving; its maintenance does so also.

All forms of material wealth wear out in course of time. Some sorts of capital are indeed very durable, such as irrigation dams and granite docks. Some last a considerable time, as buildings and machinery. Others are used up very quickly, as the coal which is burned under the boiler. All need to be replaced as time goes on: some slowly, in proportion as they last long; some quickly, in proportion as they are rapidly used up. In order that the existing apparatus of production may be maintained a certain amount of labor must steadily be given to its renewal and replacement. This labor must be supported, and its support means repeated demand upon surplus and savings.

The manner in which this takes place may be illustrated by the depreciation account which appears on the books of every manufacturing enterprise. The manufacturer knows that his machinery wears out and that, if his capital is to remain unimpaired, he must set aside something annually to replace it. Not only does his machinery wear out; in a period of rapid improvement and invention like our own it may fast become antiquated, and he must be prepared for the possibility of having to discard it even before it has ceased to be workable. If we assume that its life is ten years, he must set aside annually something like one-tenth of its value; to put it more exactly, he must put aside such sums as, invested and compounded, will make up the value at the close of the decade. If he is to secure a permanent profit, he must reckon these amounts as part of his expenses. Yet in the first instance the amounts are so much surplus, available for expenditure even tho not expected to be used for current expenses.[1] They are presumably used for purchasing new apparatus to replace that worn out but are not necessarily so used.

Commonly, capital is maintained intact; not in the sense that

[1] In practice, the actual setting aside of money, and its investment over a term of years as a separate fund toward depreciation, is probably rare. Usually, a sum is each year debited on the books against earnings, for depreciation. On the other hand, one or another item of plant is renewed or repaired each year—the whole does not become useless at one fell swoop—and the sums spent for replacement are charged against the depreciation account. In any given year, more or less may be actually so spent than is regularly set aside for depreciation. If less is spent, and the depreciation fund accumulates, it is often used, in a profitable enterprise, for putting in additional machinery or improvements—it is invested *in* the plant rather than *for* the plant.

the same machinery or materials are maintained indefinitely but in the sense that, as they wear out, other machinery and materials are regularly produced to take their place. The surpluses which are put aside to balance depreciation are again invested either in the same enterprise and the same kinds of instruments or in some other way. The habit of saving is strongly intrenched among the well-to-do. Spendthrifts are rare, and such wasting as does occur is more than balanced by the fresh accumulations of new savers and investors. Consequently the making of new capital—of materials, machinery and apparatus of all sorts—goes on constantly. The persons who in the established division of labor are engaged in the machine-making trades have a well-founded expectation that the apparatus which they produce will be bought to replace that which has worn out. The manufacturer finds new machines ready at his command. Under the division of labor provision is constantly made for anticipated needs, and among those needs that of replacing capital makes itself felt.

The repair of capital, as well as its complete replacement when worn out, calls for the recurrent exercise of saving. Some kinds of apparatus must be touched up a little from day to day in order to be in good working order. Such is the case with the roadbed of a railway, which needs almost hourly attention and would become quite unusable if neglected for a few weeks. The locomotive of a railway, again, is subjected to constant heavy strain and needs to be sent to the machine shop at frequent intervals, until finally after perhaps a generation of alternate using and patching it goes to the scrap heap and has to be replaced with a new one. The continued maintenance of capital by operations of this sort means the steady application of labor hired, thru middlemen in a successive series, by persons who mean to keep their capital intact.

§ 7. In closing this chapter something is to be said on some matters of terminology and definition. Here as elsewhere economists are perplexed by the fact that they have taken over from every-day speech some terms for quasi-technical use in their discussions, which are used familiarly in various senses with only the context to indicate just what is meant. Such terms are hoarding, saving, investment, capital. I state here in what sense they have been used

in this chapter and thruout this book, aware, however, how easy it is for any one to slip inadvertently into other meanings and connotations.

By "hoarding" is meant in these pages the putting aside of money (or jewels and the like, easily convertible into money and easily stored) with no expectation and intention of using them either for investment or for expenditure on consumer's goods. The motive may be excessive precaution against possible needs over a long period of time; or it may be the mere satisfaction from possession, a miser's gloating over the hoard, a state of mind which in older psychological speculations was sometimes classed as an "instinct" but is hardly more than an association of ideas. For the purposes of the later consideration, in Book III, of money and monetary problems, sums hoarded are to be regarded as taken out of the money stream, and for the time being as quite outside the circulating medium.

"Saving" is the putting aside of money with an eye to its use within a not distant time. The difference between saving and hoarding is only one of degree; but it is none the less an important one, and one usually indicated with clearness by the context. Some saving may be merely for convenience—the cash which one keeps in one's pocket for use at any moment, and the balance one keeps at the bank for similar early use. How fluctuating is the amount of this sort of saving by any one person, how various the length of time a given sum is kept, how changing are the habits and traditions of individuals and of groups and class—these are familiar matters. Different is the saving for eventual permanent commitment—the kind of saving that is most important as regards capital and the growth of capital. It may be for the eventual purchase of an item of consumers' capital—a house, a motor car, a piano. It may be for what is commonly called "investment," looking to a periodic continuing income over a long stretch in the future. It is the last to which I shall try to confine the word "saving."

The word that gives the greatest trouble is "investment." A dictionary meaning is the use of money—the spending of it—in the sense of acquiring something which will yield an income in the way of interest or profit. In this chapter it has been used to

mean a particular way of applying savings—expenditures that lead to the making of "real" capital, physical capital, the concrete instruments of production. It must be admitted that we commonly speak of investment in the monetary sense—the mere purchase of securities or other evidences of ownership which are expected to bring an income, with little attention or even implication as regards the meaning or outcome of such purchases for the physical equipment of the community. It is the building up of the physical instruments that has been the main theme of this chapter and is meant by "investment" as here used. Unfortunately there is no word in the language which specifically means the operation of adding to "real" capital, as indeed there is none which refers to physical capital as distinguished from capital in terms of money. Being sceptical about any advantages to be gained from newly invented terms, I have tried to use the familiar word investment to signify the process of making the physical things. It is this outcome of savings which I shall try to keep in mind when speaking of "investment."

Second, I have had in mind thruout, when speaking of physical capital, the physical instruments *made by man*,—not land or other natural resources provided by nature. Whether the word "capital" is to be applied to the latter is again a matter on which there is diversity of usage in common speech and in the writings of economists. It is best left for consideration at a later stage.[1]

[1] In Vol. II, Book V, especially Chapter 47.

CHAPTER 6

CORPORATE ORGANIZATION

§ 1. Partnerships and corporations. Limited liability. Corporations from the legal point of view and from the economic.—§ 2. Advantages from corporate organization. Large-scale operations facilitated; new and venturesome investments promoted; stimulus to savings and investment.—§ 3. Ease of transfer serves to divide risks and so promote investments, and to bring control into capable hands. But it leads to great evils.

§ 1. The growth of large-scale operations has caused a great development of combined action by promoters and managers on the one hand and by investors on the other; that is, by those who guide production and those who own the apparatus of production. Association by the manual laborers themselves for the conduct of production is a different thing. It might conceivably be an important and even dominant form of industrial organization; but in fact it is not.[1] The form which is more important than any other in the modern world is the association of capitalist owners and managers in the business corporation.

The simplest form of association by such persons is the partnership of two or more persons. The distinguishing mark of the partnership in the eye of the law was originally the joint and several liability of the partners for all debts, and this still remains in most cases. Each of the partners is liable individually and without limit for all debts of the firm. A creditor, if his claim is not met according to stipulation, may levy on any one of them and may secure the full amount of his debt from that one. The mode in which the partners then settle the distribution of the obligation among themselves is a matter with which the creditor need not concern himself.

A characteristic mark of the corporation on the other hand is limited liability. The several associated persons contribute to the undertaking, in the form of a subscription to shares or capital stock, a given sum. The liability of each for debts is then limited

[1] See what is said in Vol. II, Chapter 61, of coöperation by workingmen.

in proportion to his subscription. Usually it is limited to the precise amount subscribed. When they have once paid that sum in full—the par values of their shares—they can be called on to pay no more. Occasionally there is a different liability. For example, in our national bank corporations of former days the liability was double; the shareholder might be called on to pay not only his original subscription but (in case of need for meeting debts) as much more. Some limitation there almost always is. In the United States (almost without exception) a shareholder in a corporation is not liable, as is a partner, to the full extent of his means.

The legal distinction between a partnership and a corporation does not run parallel with that which is significant for the purposes of economic study. For the economist, the important distinction is between an association of a very few persons, well known to each other and actively engaged in the undertaking, and an association of a considerable number of persons, strangers to each other and generally investors not closely concerned with the management. Size, tho not necessarily significant, yet distinguishes roughly the two kinds of economic organization. It is true that many corporations are small, some partnerships large. But usually the conduct of operations on a considerable scale and with a considerable number of participants is in the corporate form, while partnerships usually confine themselves to more moderate undertakings.

In modern times the legislation of English-speaking countries has greatly modified the sharp distinction which the law formerly drew between the partnership and the corporation. The strict rules of the older common law made the partnership a cumbrous form of organization. It had to be wound up on the death of any partner and it was in other ways hampered in continuity of operation. Accordingly statutes have permitted partnerships to have some of the characteristics of corporations—continuing existence, inactive members, some limitation of liability. On the other hand, corporations have been allowed to enter on all sorts of industrial fields which formerly were shut to them. Originally, industrial corporations were authorized only where some special public interest was supposed to be involved; as in the case of the great companies

for foreign trade in the seventeenth and eighteenth century, of banking corporations, and in later days canals, turnpikes, railways, and the like. But the convenience of this form of associated action, compared with the cumbrousness of the partnership, caused a gradual extension of its field, until at present any and every sort of industrial enterprise may be conducted in corporate form.

The consequence is that many business corporations are of small size, owned and managed by a few individuals whose relations to each other are substantially those of partners. The choice between a corporation of this sort and a partnership of the older type is often determined by the peculiarities of the law in the place of action, by its tax methods, by its legal procedure. The fundamental distinction of limitation of liability has ceased to be of vital importance. It is true that a partnership with unlimited liability may be expected to enjoy better credit, since those who lend to it have more to fall back on. But credit in modern times depends very much on the personality and business repute of the borrowers; or, if there be question as to their business standing, it depends on the direct pledge of property. The other conveniences of corporate organization outweigh any disadvantage on the score of credit. Hence "Smith & Jones, Incorporated," or "Smith & Jones, Limited," or the "Smith & Jones Company," supersede plain "Smith & Jones"; but this change in the legal form of organization is of little economic consequence.

Very different, to repeat, is the economic significance of what we may call the true corporation. Here there are many shareholders, directors selected from among them, and managers chosen by the directors—in other words, a clear separation between owners and managers. This is the sort of organization chiefly found when production takes place on a very large scale.

In our own time, and in the United States, many people associate with the term "corporation" something still different: not only divided ownership and large-scale operations but special public importance. They think of corporations as having a monopoly power, and therefore peculiarly subject to public regulation. "Public service corporations" are spoken of as if they were *the* corporations. Whether there is a clear line of distinction between

the so-called "public" corporations and the others and whether large-scale operations in themselves bring monopoly and public responsibility will be considered in another place.[1] For the present we are concerned simply with those aspects of corporate development which have to do with the growth of large-scale production in modern times and with the modern mechanism of saving and investment. They appear not only in corporations of the "public service" kind but in others which are commonly regarded as having no special duties or relations of a public sort. In the following sections we shall consider "corporations" in the sense indicated above—those which operate on a large scale, which have many shareholders, and in which investors and managers are clearly separated.

§ 2. The advantages of the corporation for the development of industry have been great.

In the first place, large-scale operations have been facilitated. Many modern enterprises require so great a capital that no individual could supply it. In some of the older books on economics it was said that such enterprises could be undertaken only by the state and hence mere size was regarded as a criterion for public management of industry. This reason for resorting to public management can now have no force. Tho no individual or small group of individuals be able to furnish the funds needed, the corporate combination of numerous individuals can supply the means for any undertaking, however large.

Limitation of liability has been an important factor in promoting large-scale operations under corporate organization. Every enterprise involves risk, especially in its first stages. Where the enterprise is large, the amount risked and the consequent liability are correspondingly large. If each individual who took shares were liable for debts without limit, as a partner is, investment would be checked. Occasionally it has happened that a great business, conducted in essentials under corporate form but without the legal safeguard of limited liability, has met reverses and failed. Each shareholder has in such a case been subject to levy for all his property. Thus when the Glasgow Bank failed in 1878, hun-

[1] See Vol. II, Chapter 65.

dreds of small shareholders in Scotland were ruined because each was liable for the debts without limit. Probably few of them were clearly aware of this possibility when they became owners of their shares. The general practice of strict incorporation and consequent limitation of liability had put them off their guard. If experience like theirs were frequent, it would not be possible to gather the capital for large enterprises by contributions from many scattered individuals.

Further, new enterprises, both large and small and especially those which are large, have been promoted by the limitation of liability. The progress of invention in modern times, the diversification of industry, the increase of productive power—all this has taken place by successive ventures, each of which meant at the outset uncertainty and risk. It is comparatively easy to induce a person to take a few shares or even a good number of shares in a novel undertaking presenting possibilities of profit; but if participation involves also the possible loss of his entire fortune, he will be slow to join. Such a great risk will be taken only if the possibilities of profit be very great indeed, that is, if the prices of the commodity or service in question promise to be high enough to yield an exceptional profit. Limitation of liability and consequent readiness to invest in venturesome operations mean not only that more such operations will be carried on but that the community will get the output on better terms.

Probably most important of all the ways in which corporate organization has promoted the development of industry has been the ease of investment and the consequent stimulus to saving and the making of capital. In the eighteenth century almost the only possibility of investing in securities was thru the purchase of public obligations; and these, tho they meant investment by the individual, usually brought no increase in the community's capital. Merchants and persons in active business could indeed manage the investment of their surplus means in factories, warehouses, ships, and the like. But the investor pure and simple could not turn to them. If he did not buy government securities, he had little choice except to buy and improve real property. Real property is not divisible into convenient shares, and involves a good deal of man-

agement and not a little risk. The modern security market on the other hand offers an almost limitless field for the investment of savings, great and small. Railways, factories, steamships, mines—all are conducted under corporate form, and corporate obligations representing them can be bought at a moment's notice by any one. Savings have been made liquid, so to speak, and can flow with ease and in any desired volume wherever there is a prospect of their advantageous use. The ease of investment in corporate enterprise has stimulated savings; by a reciprocal influence, the unceasing accumulation of savings has made possible an immense increase of real capital under corporate management.

§ 3. The ease with which shares can be transferred from one holder to another and the consequences of easy transfer are among the most important aspects of corporate organization. This ease of transfer is by no means essential to corporate organization. Quite conceivably those who have embarked as shareholders in a company might bind themselves to stick to it for good or ill. Sometimes transferability is in fact guarded or restricted in corporations of moderate size, being allowed only with the consent of the other shareholders or with an option of purchase by them. But in the large concerns it is almost always unrestricted and is commonly thought of as a natural and necessary part of corporate organization.

Transferability, like limitation of liability, has some advantage for the community in that it makes possible a greater division of risks. A person who has invested by taking shares in a given corporation is not thereby committed to the bitter end. If he does not think well of its prospects, or comes across some opportunity which he finds more promising, he can sell his shares to another person who has a better opinion than his own of the original venture. Ease of sale in any set of business dealings facilitates venturesome operations and permits them to be carried on at a smaller margin of profit.[1] It is so with sales of securities and speculative operations on the stock exchanges. The essential advantage of such transactions for the public is that they operate as a sort of insurance against risk, and so stimulate investment, especially in new enterprises.

[1] See Chapter 11.

Transferability of shares sometimes has another advantage. It may bring ownership and control into the hands of the shrewd and competent. Those who judge best of the prospects of an enterprise and who exercise influence intelligently toward its skillful management buy out those who are less shrewd. Good judgment is perhaps the most important quality for success in business operations, and tells immensely both for an individual's money-making and for the efficient utilization of the community's labor and capital. Whether the reward which such judgment secures, often so large and so quickly won, is in proportion to the services rendered is an open question. But judgment does tell immensely for the efficient conduct of industry and transferability of corporate shares aids in making it tell.

Transferability, however, has had consequences that are clearly not beneficial. The sense of association for common ends has virtually disappeared among the shareholders of the modern corporation. Tho it persists more or less in the closely owned family corporation (the quasi-partnership), it is gone where the holders are many and scattered. Each looks out for himself and deserts the venture in case of expected loss as a rat deserts a sinking ship; or, if he expects a gain, quickly gathers in from his associates a larger number of shares for his own profit. To sell out when the affairs of a corporation are going badly, to buy in when they are going well, is the height of business acumen. This is quite inconsistent with the original notion of a joint venture for common profit or common loss; but it is not for a moment thought of as violating any principle of morals or of fair play. It may bring the advantages just mentioned: the constant buying and selling lessen risk for the individual and make for control by the shrewd and able. But it is among the phases of individualism that bring a shock to a nice moral sense.[1]

The extraordinary growth of corporate enterprises and the transferability of their shares have brought into existence the modern stock exchanges, with all their conspicuous and sometimes

[1] This disappearance of all sense of solidarity between shareholders is recognized frankly in the German practice of issuing certificates to bearer, as bonds alone are commonly issued in other countries, coupons being attached to the share certificates for such dividends as may accrue at stated dates.

demoralizing influences. The homogeneity of shares and other securities makes them available for purchase and sale by all sorts of persons and thus peculiarly adapted for speculative dealings. By far the greater part of the transactions on the exchanges have nothing to do directly with the process of actual investment; usually that has been completed before the securities are listed. It is only in the way of anticipation, thru the indirect influence of the prospect of easy transfer, that stock exchange dealings promote the increase of factories, railways, concrete capital. Tho the gain in this way is real, it is accompanied by a vast deal of unproductive effort in the way of stock gambling; nor is it easy to say whether the social gain on the whole outweighs the social loss. Most persons who discuss these matters have but hazy notions as to what constitutes the social loss or gain. They assume the corporate organization of industry as a settled fact, without discriminating wherein it is really to the general advantage. They assume transferability of shares to be a settled fact, without stopping to think whether the gain from quickened investment outweighs the material and moral loss from gambling. Still less do they consider whether the advantage which may arise from more efficient management at the hands of the shrewd outweighs the social disadvantage arising from greater inequalities in wealth.

CHAPTER 7

SOME FACTORS IN PRODUCTIVENESS

§ 1. The effect of high wages (abundant food) on the productivity of labor. High wages in the main a result, not a cause, of efficiency.—§ 2. Effects of skill and intelligence on productivity. General education. Technical education, in its effects for the individual and for the community.—§ 3. Leadership. The business man; the man of science. Freedom and mobility as promoting leadership. The motives to leadership.—§ 4. The immaterial equipment of a community; how affected by training and by inheritance.

§ 1. THE preceding chapters have dealt with such causes affecting the productiveness of industry as the division of labor, the advance of large-scale production, the use and the growth of capital. Some other factors bearing on the efficiency of labor in production will be considered in the present chapter.

Among these other factors is the quality of the laborers. The increase of production depends not only on the marshaling and organization of the laborers and on their equipment with capital but also on the strength and skill of the individual workmen. These two factors—strength and skill—may be taken up separately.

There is what may be called the steam engine theory of the efficiency of labor. It maintains, or perhaps implies rather than maintains, that the vigor of the laborer is in proportion to what he consumes. The more is turned over to him the stronger will he be and the more will he produce, just as the power got from a steam engine depends on the fuel burned in the fire box. Feed your laborer better and he will be able to produce so much more. It seems to indicate that it will always be profitable for the employer—at the least, consistent with the maintenance of profit—to pay higher wages.

There is a measure of truth in this view. It holds good particularly of the simplest unskilled labor, such as calls for continuous and heavy muscular exertion. Sometimes men are so underfed that their physical strength suffers. Employers of large gangs of laborers find that it pays to feed them abundantly. Military

operations which involve heavy labor and especially those involving long marches are more likely to succeed if the rank and file get good rations. Millions of people in backward and semi-civilized countries, such as China and India, are underfed. It is probable that their efficiency could be increased by more food and better housing. No small proportion of laborers in civilized countries are in the same situation. Mr. Rowntree, in his investigations on the city of York in England, made an estimate of the money wages which would secure, at current prices in England, the food, shelter, clothing needed for physical efficiency. The sum was about 20 shillings a week for a family of five and the earnings of one-sixth of the wage-earning class in York fell short of that sum.[1] The case is probably no less disheartening for a considerable fraction of the laborers in all parts of Europe. Notwithstanding the higher general range of wages in the United States, there are some workmen in this country also—perhaps but few relatively, yet in absolute numbers not insignificant—whose state is equally miserable.

It may seem that where laborers are underfed, an increase of wages up to the point of nourishment adequate for full physical efficiency will not be difficult to bring about, since the added product will make the added wages worth while. But the case is not so simple as it appears. Tho the laborers may gain in effectiveness from more ample subsistence, and tho the community may become thereby a healthier and happier social body, the individual who makes the advances to the laborers will not necessarily gain. If indeed the laborers were slaves there would be some chance of direct profit from feeding them better. They would remain the property of the master and he would reap where he had sown. Even as regards slaves, to be sure, it is not always profitable to go to the expense of full feeding. It may be cheaper to work them hard on poor fare, to wear them out in a few years, and to buy new ones for the same wretched round—a practice said to have been deliberately followed on some southern plantations in slavery days. Whatever the case as regards slaves, it is obvious that with free men it is essentially different. The gain in effectiveness from better fare inures to the laborer himself. Any employer who would

[1] B. S. Rowntree, *Poverty: a Study of Town Life*, Chapter IV.

make the needed advances could have no assurances of recouping himself. For the effects of full subsistence on effectiveness do not appear either with quickness or with certainty. The process is not quick, because time is needed to bring weakened and demoralized laborers into good condition; it is not certain, because some among them are so enfeebled by sustained hardship, or congenitally so weak in constitution, that they will never become able-bodied. Even tho a body of underfed laborers, if taken in hand systematically, could be brought to a pitch of full vigor, the risks and uncertainties as well as the probability that the regenerated men would betake themselves to employment elsewhere make it almost hopeless for a profit-seeking employer to carry out an operation of the kind. It is only under exceptional circumstances, where large gangs of men are at work in out-of-the-way places and are therefore under a quasi-compulsion to stick to their job—say in building the Panama Canal or at construction camps in remote regions—that it is to the immediate interest of the employer to supply the means for ample support.

The class of underfed laborers, comparatively small tho it be in modern communities, presents a distressing problem. They are ill-paid because they are inefficient; they are inefficient, for one reason, because they are ill-paid. Yet they are easily demoralized; too often they remain still inefficient if better paid from charitable funds. Neither physically nor morally do they respond readily to possibilities of improvement. Often the adults are hopeless; the children alone can be taken in hand with prospects of success. Hence even when there appears to be ground for expecting increased productiveness of labor by adding to the reward of labor, the precise method of accomplishing the result is hard to devise. Only public or quasi-public action can grapple with the problem; and this must include suppression or elimination of the unfit, as well as uplifting of the potentially capable.

All this reasoning and speculation, however, is concerned only with the minimum necessary for health and strength: the minimum, be it noted, for health and strength, not for keeping body and soul together. Men can live and do work for less than is necessary to enable them to do full work; the minimum for effi-

ciency is above the starvation level. But when they once get what is necessary for complete physical vigor, anything in addition is mere surplus; surplus in the sense that it no further increases efficiency. If obtained, it must be as the consequence of skill and productiveness; it becomes a result of high efficiency and ceases to be a cause of efficiency. Nor is the minimum for full vigor a very high one. An abundant vegetable diet, rude shelter, and simple clothing are all that a man needs in order to do the hardest work which the human frame can stand. The frugal Italian or the rice-fed Chinaman, if only he gets enough of his simple fare, can do as much as the meat-eating Irish-American.

In some of the higher walks of life the minimum for efficiency is doubtless to be measured more liberally. Something more is called for than that which is indispensable for muscular efficiency. The work of a lawyer, physician, teacher, business man, calls for alertness of mind and bodily health more than for physical vigor. The requisite response of intelligence will often be lacking if the surroundings dull the mind or enfeeble the spirit. Hence as regards intellectual work we may count varied food, ample lodging, restful relaxation among the necessaries for efficiency. It is hard to say just how far such sources of enjoyment, procured by a larger income, are really necessary for the best exertion of the mental faculties. Those who are accustomed to comfortable living and to pleasant distractions easily convince themselves that these are essential to keep them fresh for their work. It is a sort of excuse or justification of the existing inequalities in income to believe that they are necessary in the sense that the work which earns the higher income could not be accomplished without the freer life which that higher income secures. Yet plain living and high thinking are not incompatible. The luxuries and comforts to which most persons of the well-to-do classes are habituated could be in large measure foregone without loss of vigor or freshness. Some comfort, some leisure, some distraction, are doubtless necessary for the best intellectual work. But a modest income and a scale of expenditure much below that of most members of the well-to-do class would suffice.

§ 2. Different from strength are skill and intelligence. These

tell strongly on the efficiency of the individual workman and on the productivity of industry at large.

Many of the improvements in the arts depend for their application on a good degree of intelligence. The Hottentot cannot use tools even of a comparatively simple kind because his brain power is not sufficiently developed. Negroes are employed in great numbers in the gold mines and diamond mines of South Africa, but for simple pick and shovel work only; for handling and guiding the machines skilled and intelligent white mechanics must be employed. Many of the operations of agriculture require nothing beyond delving and ditching. But the fruitful agriculture of advanced peoples calls for care, discrimination, intelligence, and could not be practiced by Indian ryots, perhaps not by the peasants of backward European countries. There is a wide range of routine operations in modern industry which can be carried on by any persons capable of giving steady attention. But that very faculty, like the ability and willingness to do prolonged continuous labor, is not a matter of course. It is not possessed by savages; it is a slowly acquired quality of civilized man. No doubt there is a growing range of machine work in which very slender intellectual or moral qualities are needed. In many factory operations of modern times the human worker is hardly more than another steady and dependable automaton. Along with labor of this sort, however, there must always go some proportion of labor more flexible, more observing, more highly trained. This is the quality of mechanics' work as distinguished from that of "laborers' " in the narrower sense. Here accuracy, watchfulness, skill, intelligence are called for; and here these qualities are indispensable for efficiency.

The effect of education on the productiveness of labor is not simple. In some respects, a wide diffusion of education is conducive to greater efficiency of the population at large; in other respects, the extension of education raises economic questions not so easy to answer.

The simplest kind of pick and shovel work seems to be done as effectively by the illiterate workman as by the educated. This is also the case, as has just been remarked, with much modern factory labor. And even in many handicrafts, book education is not in-

dispensable for a high degree of skill. The work of the craftsman of the Middle Ages in Europe, and that of the same class of workers in modern Japan and indeed in some parts of contemporary Europe, show that illiteracy is no obstacle to the deftest use of tools.

Nevertheless, it remains true that a wide diffusion of education is a most effective means toward productiveness. It is effective particularly toward stimulating and diffusing new kinds of efficiency. When an art has once been learned by slow steps—and thus in the course of history mankind has acquired most of the arts—its mere transmission from generation to generation, its maintenance and even perfecting, may indeed take place by the simplest imitation, unaided by book learning. But the rapid spread and utilization of improvements are immensely promoted by the ease of intellectual communication. Mere ability to read and write opens at once a whole new world. He who possesses it can learn from the experience of all mankind, no longer from that of his parents and masters only. The extension of such a great improvement as the system of interchangeable parts has depended largely on widespread elementary education. A complex tool or machine—a plow, a reaper, a bicycle, an automobile—is made nowadays on standardized patterns, each part being a precise duplicate of every other part made from the same pattern. When there is a break, the needed part can be replaced at once. The system makes possible the wide use of intricate apparatus in localities distant from repair shops. But its adoption is possible, in turn, only if those who are to use the apparatus have some general intelligence and if they can read instructions. In the United States the unexampled use of labor-saving agricultural implements, all made with interchangeable parts, has rested not only on the intelligence of the people but on the universal diffusion of elementary education. The great industrial advance of Germany is due in large measure to the same factors.

Technical education obviously has a direct economic effect. The training of civil engineers, mechanical engineers, electrical engineers conserves from generation to generation the elaborate acquired arts. It promotes too the advance of the arts. In the past, great inventions and improvements have probably come as often

from the workshop as from the laboratory. Under the conditions of the modern world, and especially with the more methodical application of natural science to the arts, the laboratory is likely to play a larger and larger part, both directly thru the inventions that come full-fledged from the laboratory and indirectly thru the work of those who have had its training.

All training for the arts and professions tends to become more systematic in the modern world. The engineer gets his fundamental training not in the workshop or in the field but in the technological school; the physician or the lawyer gets his not from the active practitioner but from the professional school. The same movement is seen in the extension of industrial training to the familiar mechanic arts. Apprenticeship to a craftsman was for centuries the mode in which these arts were maintained and transmitted. But the conditions of modern industry have made apprenticeship ineffective and virtually obsolete. The "master" of former times has well-nigh disappeared; he is replaced by the large employer, out of touch with his individual workman whether young or old. Those preliminary stages of industrial training which were in former times provided by apprenticeship should now be undertaken by systematic trade schools and should be a part of the general system of public education. The time will come when the normal entrance to a trade will be thru such schools, precisely as the normal entrance to the so-called liberal professions is thru their professional schools.

We must distinguish sharply between the effect of such education on individuals and on the community. As between individuals, the wide diffusion of educational opportunities has simply an equalizing effect. For the community, it tends to raise general efficiency; but it is not likely to raise general efficiency in the same degree as it raises the earnings of some individuals. It tends to break down any position of mere privilege which may exist among those who now possess technical or professional skill. For these it may tend to lower earnings, while on the other hand it tends to raise the earnings of those who are enabled more easily to acquire the skill. The trade unions are usually opposed to the establishment of trade schools, from a fear that it will lower the rate of wages

in the more highly paid trades. This fear, tho much exaggerated, is not entirely without foundation. People who descant on the advantages of education and especially of industrial education often contrast the high wages of a skilled workman or trained engineer with the low wages of an unskilled laborer, and assume the difference to measure the relative productiveness of the two. They forget that if all men could easily procure the training for the better paid occupation, numbers in that occupation would be greater and pay in it would be less. Wide and free diffusion of all sorts of vocational training would almost certainly increase the productive power of the community as a whole; but it would also tend to lessen the differences in earnings which now exist and to lower the earnings of some individuals and some classes now favored.[1]

General education in all its grades, from that of the elementary school to that of the university, tho not directed to a clearly defined industrial end, doubtless has its considerable economic effects. True, it is largely an end in itself or at least a means to other ends than industrial efficiency. The mere attainment of knowledge and understanding is a satisfaction in itself, to some persons a great joy. Among man's traits none is more remarkable than his insatiable curiosity concerning all things in the heavens and the earth, and the satisfaction of that curiosity is one of the constant ends of human endeavor. But general education has its more immediate economic effects also. Tho reading and writing do not make the ditch digger stronger, and geometry and literature do not add directly to the skill of the mechanic, all education makes for intelligence, discrimination, the utilization of opportunities, the spread of improvements. It makes also for sobriety, honesty, and steady endeavor. The more it is directed to uplifting the character and training the faculties and the less it follows dull routine the more does it achieve these ends. Where it fails to achieve them, the remedy, even in the interest of bare industrial efficiency, is still not to curtail it but to improve it.

§ 3. Not least effective among the forces that bear on productiveness is leadership. It is exercised by business managers, by engineers and technical experts, and by men of science. Economic

[1] On this subject more is said below, in Vol. II, Chapter 52.

efficiency is profoundly affected by the success of a community in securing good leaders.

When intricate tools and machinery are put together by skilled mechanics, and when all this apparatus is guided to its productive outcome by still other skilled mechanics, one is tempted to say that here are the real producers. But a little consideration leads to the inclusion with them of the designers—the inventors and engineers. It requires still further reflection to include also the managers and employers. These last, the business class, seem to some persons, notably to the cruder socialists, to be mere exploiters. The real work seems to be done by the others; the business men sit by and merely levy toll. There is no greater misapprehension. The effectiveness of industry depends on the business man's leadership almost as much as that of an army depends on generalship. Under a complicated division of labor the various factors of production must be brought together and properly combined. The different kinds of labor and capital must be applied to the best natural resources. The long gap between producer and consumer must be bridged. The skilled mechanic and even the engineer would commonly be helpless without the guidance of the business leader. Especially is this the case where industry is rapidly shifting. Courage, energy, judgment, and command of capital are indispensable for economic progress. Much more will be said, as we proceed, on the significance of industrial leadership.

Another kind of leadership is that of the man of science. The progress of material civilization depends on the understanding of nature's laws. The astronomer, the physicist, the chemist, the biologist, lay the foundation for the development of the arts. Their efforts are usually stimulated in greater degree than with most men by motives of the higher sort—by the single-minded search for truth, or by love of fame rather than hope of material reward. The influence of scientific investigation on the arts, tho often indirect and unexpected, is none the less far-reaching. Faraday had no concern for the industrial possibilities when he discovered the induced current; yet how profoundly economic progress has been affected by the dynamo![1]

[1] My late colleague, Professor C. L. Jackson, called my attention to Perkin's dis-

Leaders are rare. Most men are commonplace. Among the means for promoting progress none is more important than the discovery and stimulation of those who have high abilities.

Freedom of opportunity and diffusion of education are the means for discovering those possessing unusual gifts. Among the classes of men who now lack education and are depressed by illiterate surroundings there are beyond question some and there may be many of talent, and an occasional genius. To the general advantage of a wide diffusion of education is to be added the fact that it helps to arouse and develop all the gifted. It is probable, to be sure, that high inborn capacity is most common among those to whom education and opportunity are already open. We touch here on the debatable problem of the origin and significance of social classes. There is evidence tending to show that the well-to-do are in their more favored position because they possess on the whole higher intellectual ability. But the proposition, even if established, is subject to much qualification; certainly it must be admitted that there is among the less prosperous no small fund of capacity which fails to be utilized. Tho gifted persons are probably less common in proportion to numbers among the so-called lower classes, there may be many of them. The full development in these of all their qualities for better efficiency, above all for leadership, is one of the most important objects of widely diffused education.

Freedom and democracy operate to develop to the full the scanty number of leaders. The abolition of class privileges in modern times thus has been not only of political and social consequence but has had direct economic effects also. The industrial preëminence of England during the eighteenth and nineteenth centuries was due largely to her free institutions. The lowborn person's opportunities to rise, even tho restricted, were better than on the Continent, and England profited accordingly. In the United States such opportunities have been more free than ever before

covery of purple dye, which led to the aniline dye industry, and to the investigations of Graebe and Liebermann on alizarin, which led to the manufacture of that coloring stuff from coal tar; further instances of industrial changes consequent on the discoveries of pure science. Many cases of the kind might be cited; this chapter in economic history is barely opened.

in any part of the world, and to this factor above all others is due the material prosperity of the country.

Those possessed of the qualities for leadership must not only be given a free field but they must also be stimulated to the full exercise of their gifts. Inequality of some sort appears to be indispensable as a stimulus.

Obviously we have here a question different from those considered in the preceding pages. There is an essential difference between providing a gifted person with the wherewithal to enable him to do his best and offering him a reward which will stimulate him to do his best. A reward in some way proportioned to the rarity and effectiveness of unusual faculties seems necessary to induce their exertion to the highest pitch. Such at all events has been the experience of mankind with the gift of industrial leadership. No stimulus to economic activity has yet been found comparable in efficacy to that of the prospect of large earnings. Inequality of incomes and possessions, so far as based on differences in industrial efficiency, is a potent instrument toward general efficiency in production.

This to be sure is the individualist view. It assumes that most men are influenced in their bargaining and income-earning by preponderantly selfish motives. The extreme collectivist view is that men can be readily induced to the full application of their faculties by other than selfish motives. Neither view can be maintained without qualification. As regards the leaders, some strive and work with little consideration of reward. Those having the very highest intellectual gifts in letters, in the fine arts, in pure science exercise them in pursuance of a well-nigh irresistible impulse. Industrial leadership and industrial efficiency on the other hand seem to depend on earnings and payments. Whether there are possibilities of stimulating this sort of work without inequality, or at all events without great inequality, is a question reaching into the most difficult problems of economics, and its full consideration must be postponed to a later stage.[1] Suffice it to say that under the established régime of private property material reward in the shape of high income and the chance of a fortune

[1] See Vol. II, Chapters 67 and 68.

have proved wonderfully potent and apparently indispensable in eliciting and spurring economic leadership.

§ 4. In sum, the effectiveness of industry depends not only on material equipment but also on what we may call immaterial equipment, not only on accumulated surplus in the way of capital but on accumulated moral and intellectual qualities. For this immaterial equipment maintenance and transmittal are not less important than for the community's material capital.

Education transmits from generation to generation the acquired attainments of the race, from the rudiments of reading and writing to the most elaborate technical training. Not only these intellectual attainments but moral qualities likewise must be handed down to the successive generations. Habits of industry, truthfulness, honesty, sobriety, of consideration for others, of care for the common good—all these are of slow growth, and rest on repeated example and precept.

In some degree there is transmission also by inheritance. The biologists still differ on the question whether acquired traits are inherited. The more general opinion seems to be that they are not and that only inborn qualities are passed on from parent to descendant. If this be the rule universal in nature, man also conforms to it; then some at least of the qualities that mark the civilized man can be maintained only by set training. Others perhaps have been incorporated in his nature by a process of selection —thru the weeding out, in the long course of history, of those having a less civilizable disposition. Human nature changes and improves, and the quality of men is now finer than it was thousands of years ago, perhaps than it was centuries ago. Repeatedly there are projects for hastening the process thru design—by breeding men, as animals are bred, from strains deliberately selected. Without entering here on the far-reaching questions which such proposals raise, it may be said that for the future as far as we can look into it the slow and haphazard process of unconscious selection will alone effect the transmission and possible improvement of inborn qualities. As regards the general average of ability and character, heredity leaves man from one generation to another on the whole *in statu quo*.

But persistent and repeated training not only keeps mankind *in statu quo* but also offers immediate possibilities of advance. No less than inherited quality, training contributes to make the difference between the civilized man and the savage. Man's great moral, intellectual, educational equipment must be conserved, like his material capital, by unremitting effort, and like that can be increased by effort. In both ways, the effort is largely altruistic. It results from the cares and sacrifices of parents, and from the conscious endeavor of the community to improve the quality of all its members through the diffusion of education. But it results also in no small degree from the self-regarding motives—from the desire of each individual to better his own condition and that of his family. Certain it is that man now starts from a vantage ground which makes possible still further advance. Some of his qualities for civilization he has inherited; others of these same qualities he acquires and transmits by constant effort. The outcome of all is the great immaterial capital of the community: a possession not less important for the general welfare, and perhaps not less extensible, than the capital of tools and materials.

REFERENCES ON BOOK I

On Production, see J. D. Black, *Production Organization* (1929); the only book known to me which deals well with this subject at large. It ranges into fields not covered in the preceding chapters, and can be referred to with profit in connection with later chapters, especially those in Book II and Book V.

On the division of labor and the industrial revolution, the well-ordered narrative of Mantoux, *The Industrial Revolution of the 18th Century* (English edition, 1928). On the development of inventions, A. P. Usher, *History of Mechanical Inventions,* 1929. On large scale production, vertical combination and corporate organization, the Census Monograph by W. L. Thorp on *Integration of Industrial Operations* (1924).

On corporate organization, and in general on industrial organization and control, an excellent concise account is in Tippetts and Livermore, *Business Organization and Control* (1932). A detailed account of the modern situation is in Berle and Means, *Modern Corporations and Private Property* (1932). Any book on these topics necessarily bears also on the questions of corporate management considered below in Vol. II, Book V, Chapter 50.

BOOK II

VALUE AND EXCHANGE

CHAPTER 8

INTRODUCTORY: EXCHANGE, VALUE, PRICE

§ 1. Exchange the consequence of the division of labor.—§ 2. Money as the medium of exchange.—§ 3. Value and utility. The notion of value in exchange.—§ 4. A general rise in values, a general rise in prices. Stability in general prices provisionally assumed.

§ 1. THE division of labor brings in its train the exchange of goods between those who undertake the separated acts of production. Exchange in turn brings the phenomena of value, money, and prices. With these phenomena we shall be concerned in the present Book and in the Book following.

As has already been noted, the division of labor does not bring exchange as a necessary consequence.[1] There may be the self-sufficing patriarchal family, with a division of labor but without exchange; or its counterpart, the small communistic society, self-sufficing at least to some degree. Even in the modern family there is division of labor, after a sort, between man and wife. But commonly we consider the family as a unit, and think of the housewife, when she works for husband and family, as working for that of which she is but a part. Similarly, the patriarchal family and the communistic society are regarded by their members as social and economic units. Exchange arises from a separation of interests and has grown with the growth of private property.

Production for one's self holds its own longest in agriculture. Yet even in this industry division of labor and exchange are rapidly extending. In the United States the self-sufficing farmer of earlier days has well-nigh disappeared, and even the stolid peasant of Europe is being transformed by the modern methods of easy communication and ready sale and purchase. Tho the farmer still produces part of his own food, especially vegetables and fruit, there is a steady tendency toward widening the range of agricultural products which are bought and sold. Grain is sold by the

[1] See Chapter 3, § 4.

individual farmer, flour is bought; cattle and fowl are sold, meat is bought; milk and cream are sold, butter is bought. In other occupations the division of labor has worked out its consequences to the last stage. No labor is given to the direct satisfaction of each worker's wants; all is turned to the indirect process of specialization and exchange. Sale, price, value, and the whole mechanism of exchange become the characteristic economic phenomena.

§ 2. Almost as early as the division of labor a medium for exchanging various products came into use. Barter—the direct exchange of products—may be carried on under a very simple division of labor; yet even then it is inconvenient. As soon as the first stages of savagery have been passed some use of a medium of exchange appears.

Any commodity which is in general use will serve passably as a medium of exchange. He who has an article to sell, and cannot find at once the precise kind and quantity of the things he wishes to buy, will accept a staple commodity with which sooner or later he will be able to procure the things he wants. Hence in various stages of civilization the most diverse commodities have been used to obviate the inconveniences of barter. In Homeric times the value of things was often stated in terms of oxen, for such occasional exchanges as are made among primitive pastoral peoples are naturally effected in terms of their staple commodity, cattle.[1] For a considerable time in the early history of the colony of Virginia, tobacco was almost the sole article of export, and the chief commodity habitually produced for a market; it became the recognized medium of exchange in the colony. Furs, salt, tea, cocoa, have served the purpose with other people. But by far the most widespread among the things so used have been the precious metals, gold and silver. We need not pause at this stage to consider what qualities fit them peculiarly for serving as a medium of

[1] Mr. Wicksteed remarks (*The Common Sense of Political Economy*, p. 137) that "there is more evidence in the Homeric poems of the valuation of female slaves, of tripods, or of gold or brass armor, in terms of so many cattle, than there is of any direct transfer of cattle in payment of those goods." It is probably true of the other commodities mentioned in the text that they also were used more freely for measuring relative values than for effecting exchanges.

exchange—their luster and consequent attractiveness for ornament, their freedom from rust and deterioration, their homogeneity, their divisibility. Nor need we consider how the device of coining has increased their fitness for carrying on purchases and sales; nor in what ways paper representatives or substitutes for them have come to be so widely used in our own time. These are topics that belong to the subject of money, to which attention is given in the next Book.

It suffices here to note how completely division of labor and exchange work out their results thru the use of money. Every producer gets his return in amounts of money. The exceptions in any of the countries of advanced civilization are so few and are so rapidly disappearing that they serve only to make clear how virtually universal is the rule. Exchange takes place by first selling goods or services for money, and by then buying with the money such other goods and services as are wanted. The fundamental fact of exchange is thus obscured by the very mechanism that so perfectly facilitates it. Just as the coöperation and combination which are essential features of the division of labor are carried on without a consciousness of any combined action, so the process of exchange goes on without the consciousness that it is the aim and end of all buying and selling.

§ 3. The value of a commodity means in economics its power of securing other commodities in exchange. It means the rate at which the commodity exchanges for others. If a bushel of wheat can be exchanged for a large quantity of other things—for many pounds of iron, many yards of cloth, many ounces of salt—its value is high; the possessor of it can procure many of these things. If a bushel of wheat can be exchanged for but few pounds of iron, few yards of cloth, few ounces of salt, its value is low; the possessor of it can procure few of these things. It is immaterial that the exchange does not take place directly, but by the process of first disposing of the wheat for the medium of exchange—money—and then procuring with the money the iron, cloth, salt or other desired commodities. The result of the double operation is the same as if the exchange had taken place by direct barter. Only it is reached by a more convenient method.

The value of a commodity, thus conceived, is its value in exchange. This is very different from its usefulness, or utility, or importance. In everyday discourse we use the word "value" sometimes to indicate exchange value, sometimes to indicate utility or importance. We speak of the value of iron as greater than that of gold, and the value of wheat as greater than that of diamonds. We mean thereby that iron and wheat are more important, satisfy more urgent wants than gold and diamonds. Yet we also speak of gold and diamonds as more valuable commodities than iron and wheat; then we use the terms "value" and "valuable" in the sense of value in exchange, and mean that exchange and sale take place on such terms that with comparatively little gold and diamonds the owner can secure much iron and wheat. For the purposes of economics this latter sense, exchange value, is the more important.

A third sense, however, may be noted in passing. People sometimes speak of the "value" of a thing as greater or less than that which appears in an actual transaction of exchange. They speak of a house as being "worth" more than they paid for it, or of a commodity or a stock exchange security as selling for less than its "intrinsic value." They mean that the current price is different from the price that is likely to be paid in the long run, or different from the price which they think proper and just. In the sense which we have adopted, value means simply the actual rate in exchange, and there can be no value other than that registered by sales and exchanges. That the word is also used with this third signification, of proper or eventual worth, only shows how uncertain is everyday phraseology. Economists have often pointed out how much troubled they are by having to employ familiar terms, like capital and value, which in everyday use have various and shifting meanings. For the purposes of economics one meaning or definition should be selected and held to with care. If there be a different meaning, this should be clear from the context, if not made so by specific explanation. In the following pages I shall use "value" in the sense which economists have adopted for it—a relation in exchange.

By the *price* of a commodity is signified the amount of money which it will command; in other words, its value in terms of the

accepted medium of exchange. That accepted medium of exchange and money of account was for centuries a coin containing a given amount of one of the precious metals. Prices were stated and measured in terms of such coins—silver or gold "pounds," silver or gold "livres" or francs, silver or gold dollars. These prices indicated and measured the values of goods. The commodity which sold for a high price had high value in exchange, as compared with one selling for a low price. In the course of the nineteenth century there came a great complication in the make-up of the money in terms of which prices were stated and values measured. Gold displaced silver as the chief money metal, and an enormous increase took place in the use of paper substitutes for metal. In the first three decades of the twentieth century further momentous changes set in both as regards the use of gold itself and that of the paper substitutes. The meaning and consequences of these changes will receive attention at a later stage. For the present we may speak simply of the pounds sterling or francs or dollars with which all the world is familiar. The money paid out as we make our daily purchases is the price of the goods bought; in everyday usage this is reckoned to indicate their value, a usage which it will be convenient to follow for the present.

§ 4. From our definition of value it follows that there can be no general rise in values and no general fall in values. Value is a term expressing the relation of exchange between commodities. If at a given time a commodity procures in exchange less of others than at an earlier time, it has fallen in value; but *pro tanto* those other commodities have risen in value. All cannot rise and fall together. A change in the value of any one of them, or any set of them, means a converse change in the value of the rest. On the other hand, a change in general prices is not only possible but is one of the familiar and recurring phenomena of economics. Wheat, iron, diamonds, things in general, may all exchange for more dollars now than they did ten years ago; and ten years hence they may exchange for less dollars than they command now.

Evidently a general rise or fall in prices signifies a change in the value of money. When all prices rise, and things exchange each for a greater number of dollars, the dollar can buy less than

it did. Its power of commanding other things is less and its value has fallen. When every single thing exchanges for a smaller number of dollars, that is, when prices have fallen, the dollar buys more and its value has risen. The value of money is inverse to the level of prices. When prices are high, the value of money is low; when prices are low, the value of money is high.

The mere fact of a rise or fall in the price of a single commodity, therefore, does not indicate whether or no its value has changed. It may be that other commodities likewise have fallen in price to the same extent. In that case there has been a rise in the value of money and a fall in the value of all commodities as compared with gold, but no other change has occurred in the values of commodities. Or it may be that this single commodity alone has fallen in price, others remaining as before. In that case the fall in its price registers also a fall in value.

The general level of prices changes but slowly. While prices of individual commodities change quickly, all do not as a rule change quickly in the same direction. A rise in the price of any one is likely to be accompanied by a declining price of another, or by stationary prices of the others. So gradual are changes in the general range of prices, so uncertain the comparison and offsetting of the complex individual changes—a rise here, a fall there, no change at all elsewhere—that it is often difficult to ascertain whether during a short period a general change has taken place at all. If, indeed, an upward or downward movement continues for years, it becomes unmistakable. We can make sure then whether the value of money has risen or fallen, and can measure with some accuracy the extent of the change. But unless the lapse of time exceeds two or three years, it is often not easy to determine what has been the general trend; so stable are prices for short periods.

This stability of the general level of prices, it should be remarked, is by no means universally maintained. True, so long as the medium of exchange consists of gold and of other forms of money convertible into gold, there is ordinarily no occasion for qualifying what has just been said. Yet even with a currency resting on gold, great and rapid changes affecting all commodities

may take place in a short period of time—within a year or two; as was shown by the sharp rise of all prices in the United States during the war of 1914–18. Changes even more violent may occur when the gold basis has been abandoned and resort has been made to paper money pure and simple. These, however, are monetary phenomena to be considered at a later stage. It is of the first importance to keep clear the distinction between the forces that act at any given time on all prices—monetary forces we may call them —and the more complex forces that determine the price of one commodity compared with another. It is the latter that will be considered in the chapters to follow.

Tho general prices and the value of money as a rule change but slowly, the prices of individual articles change quickly and considerably. The price of wool or iron may rise by ten per cent in the course of a month; and changes are common in the prices of individual articles—of wheat, cotton, copper, coal—by ten, twenty, fifty per cent in the course of a single year. Where the price of one thing changes, other prices remaining the same, the new price evidently registers a change in value. The ordinary fluctuations in the prices of things thus signify, to repeat, corresponding changes in their values.

For the purposes of an orderly and systematic exposition of economic principles we shall for the present assume stability in general prices; hence that a change in price of an article signifies a change in its value. If an individual article rises in price under these conditions, it commands more of other things in exchange and rises in value, and conversely if it falls in price. The familiar examples of price and money will be used in the illustrations and figures, the problems of fluctuations in the general level of prices being put aside for consideration at a later stage.

CHAPTER 9

VALUE AND UTILITY

§ 1. Utility a necessary condition of value; but value not proportional to utility.—§ 2. Increase of supply brings lowering of value; because of differences of means, and, fundamentally, because of the tendency to diminishing utility. Effects of varying the commodities supplied. Possible exceptions to the general principle.—§ 3. Total utility and marginal utility.—§ 4. Value depends on marginal utility. Qualifications and explanations. Marginal vendibility. The marginal utility of money.—§ 5. Consumer's surplus. Sundry limitations on its significance and on the possibility of measuring it.—§ 6. How state and measure the income of a community? —§ 7. The law of diminishing utility points to the conclusion that inequality lessens maximum well-being.—§ 8. Concluding remarks on welfare and total utility.

§ 1. An object can have no value unless it has "utility." That word is here used in a special sense. Like "value" and other words which economists have come to use, it has in common speech various meanings and shades of meaning. In the discussion of exchange and value which follows it refers to the satisfaction or gratification derived from an article at the time it is procured—its filling a want or a desire, having the quality of desiredness. The terminology is perhaps not the best, but is sufficiently precise and is now so imbedded in the literature of the subject that it is convenient to accept it.

Distinctions are to be noted between the satisfactions that are expected and those that are realized, and also between the immediate and the enduring. People are sometimes short-sighted and buy things as children do, to please a moment's fancy. At the moment when they part with something—money—in order to procure an article their conduct shows merely that gratification is expected, not that it is secured. Again, people often buy things which, while they gratify at the moment, in the end lead to harmful and painful results, such as strong drink or narcotics. In all these feelings and responses it is the decision of the purchaser at the moment that tells. What counts is that men are willing to give up something in order to procure things that gratify, that promise to give utility;

this is what determines the price they pay at any given stage. At another stage they may act differently with regard to identically the same articles, but it is always the utility expected at the time of purchase that actuates them.

But the utility or gratification from a thing is not necessarily, perhaps not usually, in proportion to its value (i.e. exchange value or price). Desiredness is an indispensable condition of value, but does not necessarily make an article valuable, or determine how great is its value. Fresh air, pure water, the beauty of nature may gratify immensely yet be without value. Other things of great utility may have but slight value. The simplest and most welcome articles of food have low value in advanced communities, tho they satisfy the most pressing wants. So it is often as regards clothing, shelter, warmth. High utility often goes with low value. Other things whose exchange value is high have utilities which we do not ordinarily reckon great. Tasteless ornaments, a stupid book printed four hundred years ago—such things sometimes command a high price. They are much prized by some persons, tho others may wonder what sort of satisfaction they can yield.

§ 2. The supply of a commodity, as we all know, closely affects its value. If at any given time an article becomes more abundant, its price falls; if supply becomes lower, its price rises. The causes of these fluctuations are two, very different in nature and in social significance.

One obvious cause, and that which many persons are likely to think of first, is the difference between rich and poor. Those who are able to pay highest secure the first installments of any commodity that comes to market. If there be comparatively few installments, each will command a high price. As more and more are offered, the price must be lowered in order to bring them within the means of the less rich. If the supply be greatly increased, the price must be lowered very much in order to make purchases possible by the poor.

But the same result, tho not the same gradation, would appear if there were no differences between rich and poor—if all persons had the same incomes. Then also an increasing supply would bring a decreasing price. The principle which explains why the same

inverse variation would appear under equality of incomes is that of diminishing utility; and this second cause is the more fundamental, since it underlies the first.

Consider any familiar article of daily use—the knives, forks, spoons on your table, the clothing you wear, the house you live in. One set of knives and forks is essential to cleanly eating. A second set is highly convenient, a third somewhat less so; there is a decline in utility, until at last the stage is reached where an additional set is a mere encumbrance. So with clothing. One suit is necessary; a second and a third add to comfort. More and more can be used, yet with a steady tendency to lessening satisfaction from the successive installments. One room in a house, or a one-room house, may be indispensable for existence. The added comfort and decency from a second room are very great, and further additions to the houseroom continue to yield satisfactions. The rate of diminution in utility may be for some time comparatively slow but the tendency still is present, and before long the stage is reached when more houseroom serves to satisfy only the love of display, not to yield substantial comfort. All things, it may be observed, which minister to the love of display, have the possibility of maintaining this sort of utility to a curious degree. The mere fact that a thing is rare—that it is of a sort not possessed by others, and so distinguishes its owner—gives utility to things otherwise worthless; a notable example is an old postage stamp. Additions to the supply of many classes of articles may for a long time give additional satisfaction, if the individual things be varied and adapted to gratify the love of distinction—as with the garments and houses of the rich. But the tendency to diminishing utility none the less persists. The addition of a new coat to an abundant supply, of a new room to a house already large, brings less satisfaction than the preceding coats or rooms brought.

To this general tendency we give the name of the law, or principle, of diminishing utility. Successive doses of the same commodity or service yield diminishing utilities. If the doses be continued indefinitely, the point of satiety will be reached. Their further repetition yields no satisfaction whatever; the utility of each additional dose becomes *nil*. This principle, as has just been

intimated and as will presently be explained further, applies in strictness only to units of the *same* commodity (or service). Vary the thing supplied—even tho it be made unlike only in small degree—then the result will be different. The diminution in utility may be prevented or checked, and the point of satiety may be indefinitely postponed.

To put the proposition in other terms: all enjoyments tend to pall if repeated. If any one of us were called on to retrench—to dispense with some enjoyments now possessed—he would find himself cutting off first those things least prized, and then in succession various others in the inverse order of their utility. The process would make it clear not only that some things have more utility than others but that some doses of the same thing have more utility than other doses of that thing.

It is this fact of diminishing utility that explains the growing variety in the articles produced and the growing complexity of production and consumption. As the productive power of mankind increases, and especially as the commodities in common use become more abundant, labor is constantly turned in new directions. It is given not so much to making more of the same things as to making different things. Abundance without variety means that the approach to satiety is rapid. Bread, in most civilized countries, is cheap, being produced with comparatively little labor. With increase in the effectiveness of industry, more and more bread could be produced with the same labor. As bread becomes cheaper some of this labor is turned to other kinds of food—to meat, eggs, butter, vegetables, fruit, sugar. A varied diet, such as is possible in modern times, marks a great advance not only over the monotony of savages' food but over the very restricted diet with which civilized peoples had to content themselves until the last century or two. The essentials of clothing also are plentiful and cheap, and a comparatively small part of the labor of a modern country is given to the covering needed simply for health and decency. A vast deal of labor is given to making more convenient and attractive clothing. Variety in production must take place if consumption is to respond.

There are articles to which the principle of diminishing utility

does not apply as unfailingly as the preceding statement suggests. Tho stimulants on the whole show unquestionably the tendency to lessening response, the response to the first few doses does not always indicate it. The second or third glass of liquor may be as much enjoyed as the first. Or, to speak of higher things, the second or third reading of noble verse, or hearing of beautiful music, often gives greater pleasure than the first. There are many cases where a preliminary stage of doubtful satisfaction is succeeded, with habituation, by unquestionably greater satisfaction—as with tobacco and oysters. Many a novel article needs to be insinuated into people's liking. As this is brought about (perhaps by skillful advertising) the article reaches a stage where a larger supply of it is sold, not at a lower price per unit but at a higher. In such cases, however, the tastes of the purchasers have changed in the interval; at any given stage of taste and popularity the principle of diminishing utility will apply.

§ 3. From the law of diminishing utility we are led to the conceptions of total utility and of marginal utility.

Utility, in the sense in which we are using the term, can be measured in one way only: by the amount which a person will give to procure an article or a service. Enjoyment or satisfaction is subjective. The objective test of it is willingness to pay. What a person stands ready to give for an article rather than go without it is the only test by which we can ascertain with any approach to precision how much satisfaction it brings him. Price, actual or potential, is the measure of utility. Not infrequently in discussion of this set of subjects it is said or implied that the utility of an article *is* the price it commands or might command. The language is inaccurate. Price simply *indicates* utility.

Consider now how price may measure the utility to an individual of several units of a given commodity—say five oranges. Suppose them to be offered in succession, each being appraised by itself without thought of there being more to come. The first we may believe to be so fragrant and refreshing that the individual would pay 50 cents rather than go without it. The second, yielding less satisfaction, would command only 25 cents; the third would command still less, say 15 cents; the fourth, 10; and the fifth, only

5. The total utility of the five would be indicated by the sum of the amounts which the several units would have commanded separately, namely:—

For the first orange	50 cents
For the second orange	25 cents
For the third orange	15 cents
For the fourth orange	10 cents
For the fifth orange	5 cents
For the total supply	105 cents

Suppose now, on the other hand, that the five oranges exist as a stock, possessed together by the individual. All are alike. Take away any one and the loss of utility or satisfaction will be the same as if any other had been taken away. Each has the same degree of importance for his welfare. As installments or successive doses they have differing utility, but possessed as a stock they have each the same economic importance. Any one would be parted with on the same terms as another. And those terms—the price—would be settled by the utility (satisfaction) yielded by the *last* of them if they were enjoyed in succession. The price at which the fifth would be bought (or sold) is the price at which any one of a stock of five would be bought (or sold). That price measures the *marginal* utility, or *final* utility, of the supply to this individual. Economic importance, marginal utility, final utility, the satisfaction got from any one unit of a stock—all these expressions come to the same thing.

It may seem paradoxical to say that all the constituents of a stock have the same economic importance and that none the less some have greater utility than others. But there is no real paradox. It must be remembered that utility means satisfactions or enjoyments. To possess a stock is not to enjoy it (except so far as, by association of ideas, mere ownership gives pleasure, as in case of a miser's hoard). The stock is necessarily enjoyed not as a whole but by installments; and as it comes to be so enjoyed the successive installments yield lessening satisfaction. Economic importance is something different: it is the satisfaction dependent not on the whole stock but on any one of the constituents of the stock. Considered in this way, all the constituents are alike; even tho, con-

sidered as sources of enjoyment when actually used, they are of varying utility.

§ 4. Let us return now to the relation between the supply of an article and its price. In doing so we pass from a consideration of the individual's satisfactions to those of a group of individuals, and thereby are brought to a consideration not only of marginal utility but also of marginal vendibility.

Increase in supply means lower price. It also means lessening utility from the added units. The price of a commodity depends, as the case is commonly stated, on the least of the utilities yielded by the supply, or on final utility; price, or value, depends on the utility of the last increment. Properly qualified and properly understood the principle is sound, and not only so but of primary importance.

First as to the qualifications. The proposition is true, in strictness, only if we suppose many competing buyers and sellers. And in fact most things are brought to market by competing sellers, and are purchased by competing buyers. Assume now that a given supply, say 1000 oranges, is offered by the sellers. Among the buyers are some whose means are large, others who value oranges highly. Both sets would be willing to pay a high price for a few oranges rather than go without. But there are more oranges than these purchasers are eager for. To induce the rest to buy, or to induce the eager purchasers to buy more, the price must be lowered. As the sellers are many and competing, the price of the whole supply will be uniform. Any one seller, trying to obtain a higher price from the eager buyers, would be undersold by others. There would be one price at which the whole lot would go, and that would be the price which tempted the last buyer; or, to be more accurate, the last purchase by any of the buyers. This last purchase, and the price which must be offered to induce it, would settle the terms for all the transactions.

Next, as to the just understanding of the proposition. Observe that the last buyer and the last purchase have been spoken of, not the last or marginal utility. In statements of this fundamental principle of value it is often said simply that selling price, or exchange value, depends on marginal utility. The assumption is

here tacitly made that all the buyers are in the same position and that all have the same means. From this it would follow that a less sum of money paid out denoted a less satisfaction, and that he who bought the last unit of the whole supply was not only the last purchaser but the purchaser to whom that unit gave the least satisfaction. The fact is, however, that purchasers have very different means and, as just pointed out, this circumstance is of vast importance in explaining the fall in price which actually takes place when supply is increased. The dependence of selling price on the last purchaser (or the last purchase) is not affected by the inequality of incomes. But the significance of the final purchase for the utility or satisfaction-yielding power of the last installment of the supply is much affected.

We may speak, therefore, of marginal *vendibility*. The common formulation by economists, that price depends on marginal utility, tacitly ignores the effects of inequality. The term "vendibility" points to the dominant position of the price at which the last item is sold, and makes no implication concerning the satisfactions secured by the person who pays this price. Marginal vendibility is the resultant of two forces, diminishing utility of successive units and inequality of incomes. So far as concerns the immediate determination of price and the mechanism of supply and demand, it is not material which of the two happens in a given situation to be of most effect. The outcome is the same: increase of supply leads to a decline in price. But the social significance of price fluctuations and of the working of supply and demand is very different according as the one or the other is of controlling influence.

The simple and familiar fact that a rich man, when paying out a given sum of money, often gets less satisfaction than a poor man, when he pays out the same sum, is expressed in more technical terms by saying that the marginal utility of money is less to the rich than to the poor. A dollar signifies little to the man of means. If he parts with it, his loss in welfare is vastly less than that of the poor man who parts with the same amount. A high price therefore does not necessarily indicate great utility to those paying it. It may signify only that the marginal utility of money is small.

The phrase "marginal utility of money" must, however, be used

with caution. Money has utility in a different way from other things. It is valued not because it serves in itself to satisfy wants but as a medium of exchange having purchasing power over other things. Gold jewelry is subject to the law of diminishing utility precisely as other things are. But gold coin—money—is subject to it only in the sense that an individual buys first the things he prizes most, and then other things in the order of their smaller utility. Strictly speaking, the statement that money has varying utility and that there is a marginal utility of money is only a way of saying that the things bought with money have varying utility, and that some among them are at the margin of utility.[1]

§ 5. The conceptions of total utility and marginal utility lead to that of consumer's surplus.

Consumer's surplus is the phrase applied by Professor Marshall to the difference between the sum which measures total utility and that which measures total exchange value. The total exchange value of a set of goods is obviously the price per unit multiplied by the number of units. But the total utility of the units as they come to be enjoyed is a different quantity. Thus, our orange-eater would have been willing to pay for the first orange 50 cents, but had to pay only 5 cents. He had a "surplus" of 45 cents' worth of satisfaction. Using the same figures as before for the supposed supply of five oranges, we get the following comparison between the prices that would have been paid and the prices that were paid in fact, the difference indicating consumer's surplus.

	Potential Price, Measuring Total Utility	Actual Price	Consumer's Surplus
For the first orange . . .	50	5	45
For the second orange . .	25	5	20
For the third orange . . .	15	5	10
For the fourth orange . .	10	5	5
For the fifth orange . . .	5	5	
For the whole supply . . .	105	25	80

[1] See what is said further on this topic, and on the peculiarities of the value of money, in Chapter 18.

The case is stated here in the simplest terms, and on the assumption that the price of this small supply of oranges would be determined as is the price of the usual large supply of commodities as they come to market in the actual world—by the price which carries off the last increment. Without stopping now to inquire how far this assumption in fact holds good where a very few commodities are put on sale,[1] let us consider the nature of consumer's surplus as here typified.

How substantial is this surplus? and how far is this mode of measuring it satisfactory? To ask these questions is only to ask, in different words, how substantial total utility is and how far we can measure total utility.

One limitation of the first importance has already been indicated when considering marginal utility and its connection with demand. If all persons had the same income, then willingness to pay a given amount for an article might be fairly assumed to mean that the article had the same utility for each of them. But some have greater incomes than others; the marginal utility of money is less to the rich; and the payment by them of a larger sum does not signify a higher utility. Price depends—to use the phraseology suggested a moment ago—on marginal vendibility, not simply on marginal utility. A rich man will pay for hothouse fruits or vegetables a sum quite out of the question for a person of modest means. The latter might secure at a season of greater plenty precisely the same things for a price much lower. The rich man probably gets no greater enjoyment from his expensive purchase; or, if so (counting as part of his pleasure the gratification of the love of distinction), by no means in proportion to the higher price he pays. If some of the familiar comforts of civilized life became scarce—fresh milk or good bread—they would command a high price, even if all persons had the same incomes. But the price would go still higher if there were a circle of persons able and ready to bid heavily for them without making serious gaps in their incomes. The special increase of price resulting from this latter circumstance is indicative not of specially high utility but of large means for purchasing utilities.

[1] See Chapter 10, § 9.

Still another qualification is suggested by this fact of inequality. Many articles which command a high price satisfy the passion for display. Such are the precious stones, rare paintings and statues. No doubt many things of this sort—the great works of art—are intrinsically beautiful and yield enduring and unalloyed pleasures; and it is their intrinsic beauty, tested by time, that is the basis of their high value. Yet since they are rare as well as beautiful they satisfy also the deep-rooted instinct of emulation and desire for distinction. They have what has been called a prestige value, commanding a higher price simply because they are already high in price. Suppose now that such things became common and therefore cheap, that diamonds, for example, became very plentiful and that their price fell to some such level as that of glass beads. The intrinsic qualities of diamonds would remain: their luster and brilliancy, their hardness. The satisfaction which the previous limited supply had given might be thought, therefore, to remain undiminished. Yet in fact it would be vastly diminished; for diamonds would no longer be evidences of wealth and social station. Consumer's surplus, as measured by the previous high price, would evaporate.

Consumer's surplus is thus unsubstantial for a considerable range of articles much esteemed and paid for at high prices. Not only the favorite objects of rich collectors, such as rare paintings and books, belong in this class but many others which are not commonly thought of as belonging there. Handsome houses, fashionable clothes, even choice food, get no small part of their power of yielding utilities from their satisfying the sense of distinction. As to all these, total utility and consumer's surplus are highly elusive.

Another qualification concerns articles at the other end of the scale—things of simple necessity. Measured in terms of the prices that would be given for the early doses, consumer's surplus is very high for bread, clothing, houseroom—for the minimum of food, raiment, and shelter. Rather than dispense with these, anything would be given; life itself depends on them. Total utility and consumer's rent might be calculated to be infinite. Certain it is that were these things to become scarce, their price would go to a very high range; and this irrespective of whether there were or were

not inequalities of incomes among the purchasers. But a question may be raised as to the nature of the utilities derived from bare necessaries. The satisfaction they give is of a negative sort. The chronicler of Lewis and Clark's expedition across the American continent narrates that at one stage the explorers subsisted on dried salmon in the form of a tasteless powder, so unappetizing that only the absolutely necessary amount was eaten. Some such situation was in the mind of an ingenious and stimulating thinker, Professor Patten, who distinguished between a "pain economy" and a "pleasure economy." The first phrase describes that economic stage in which the efforts of man suffice only to yield the indispensable minimum: to prevent hunger, thirst, freezing; to ward off pain, not to yield satisfaction. The second describes that better stage when the first elemental wants have been attended to and positive enjoyment begins: when food is appetizing as well as sufficient, when clothing and houseroom are attractive. Now in reckoning total utility and consumer's surplus we do well to begin only when this second stage has been reached. Let those utilities which are of the indispensable sort be set aside. Only where the stage has been reached of possible comfort, of some choice in the direction of expenditure, can there be anything in the nature of a real surplus of satisfaction for the consumer.

This is true not only of absolute necessaries but in a good degree of conventional necessaries. Equipages and horses were formerly conventional necessaries for many members of the Continental aristocracy. They were immensely missed if the individual had to give them up. Yet the real enjoyment from them was doubtful. So it is with the starched linen and close-fitting clothes of the well-to-do, which are insignia of the wearer's exemption from manual labor. The satisfaction from them is chiefly negative: their loss would be more keenly felt than their presence is enjoyed. Positive satisfaction is indicated in very uncertain degree by the price which under the stress of convention the individual would pay for such things rather than do without.

Not the least of the difficulties in the way of measuring utilities by potential prices is the practical one that we have no means of knowing what prices would be paid for the several installments

of a commodity if they were offered one by one. In our illustrative case it has been assumed that the first orange would be so greatly enjoyed as to command a price of 50 cents. But in hardly any actual case do we know what price would have been fetched by the first installment or by a series of earlier installments. All we know is that they would command much more than that settled by marginal vendibility for the actual supply. We have some information (tho not very much even here) regarding the variations of prices in the neighborhood of the range familiar to us. We observe how oranges, cigars, bread, meat, sugar, go up and down as the quantities become somewhat greater or less than those usually put on the market. But we have no precise knowledge of what would happen if the quantity were to vary greatly from the usual amount. Statistics of prices, however perfected, throw no light on the very high range that would be paid if supply became very small.

These accumulated difficulties make it impossible to measure in any precise way total utility or consumer's surplus. The figures which have been given for illustration are useful in making the conception clear but are misleading in that they imply accuracy of measurement. We cannot set down the complete price schedule; and even if we could the differences in incomes, the illusiveness of prestige, the doubtful satisfaction of a pain economy, combine to render a calculation of real enjoyment impracticable. We cannot measure with any approach to accuracy the satisfactions got from wealth.

None the less, total utility and consumer's surplus are not fanciful. That they are real is shown by their accord with familiar phrases. We often say that we get a thing for less than it is worth to us, meaning that what we give for it offers less satisfaction than the thing we buy. This is merely stated with more care and precision when we say that a consumer's surplus is secured. Tho that surplus may not be measurable with accuracy either at the lower end of the scale of consumption where bare necessaries alone are bought, or at the upper end where mere vanity is satisfied, it is unmistakable in regard to what may be called the enduring comforts of life. A varied diet, abundant houseroom, clothing and fit-

tings that permanently please the taste, the gratification which all get from the mimic arts, distraction coming after monotonous work, the pleasures of the intellect—these are things not less enjoyed when abundant and cheap. They often have a utility much greater than is indicated by the price of them. Tho their utility be not susceptible of measurement, total utility is large and consumer's surplus is correspondingly large.

§ 6. The discussion of utility, total utility, and consumer's surplus leads to another question: how state and measure the income of a community?

An individual usually thinks of his income, and measures it, in terms of money. So long as the prices of commodities and services remain the same this mode of estimating income is for most purposes sufficient. The condition stated—of stable prices—is obviously important. If all money incomes double, and all prices also double, the community is no better off than before. It simply conducts its exchanges with a different scale for the medium of exchange.

Hence we proceed naturally to the next step. Money income is significant simply as a way of measuring the quantity of the things which the money buys. We may think, therefore, of real incomes in contrast to money income—of the necessities, conveniences, and luxuries of life. We must reckon also, as part of real income, the services of those who used to be called "unproductive"—actors, musicians, servants, and so on. The more we can get of such "real" income, of all kinds, the more prosperous we are as individuals and as a community.

But we may go a step beyond. We have seen[1] that production consists in the creation of utilities. Now just as all production in the last analysis consists in the creation of utilities, so all income consists in the utilities or satisfactions created. Economic goods are not ends in themselves but means to the end of satisfying wants. In a preceding chapter we have distinguished between capital and wealth which is not capital, or (in other phraseology) between consumer's wealth and producer's capital. But consumer's wealth, which we may treat in one sense as "real" income, is an

[1] See Chapter 2, § 2.

instrument no less than producer's capital. It too is a means, not an end. Our food, clothing, furniture, may be said to yield psychic income. They shed utilities, so to speak, as long as they last. In the final analysis the income of an individual or of a community consists of the sum of utilities steadily accruing from its store of goods and services. It consists, that is, of the total utility of all.

Nevertheless, for almost all purposes of economic study, it is best to content ourselves with a statement, and an attempt at measurement, in terms not of utility but of money income or of real income. The reason for this rejection of a principle which is in itself sound lies in the conclusion just reached regarding total utility and consumer's surplus: they cannot be measured.

The other ways of stating and measuring income lead to results of some certainty. We can measure money income. Tho our statistics for the total money income of (say) the people of the United States are far from complete, the task of ascertaining that income is not hopeless. Indeed, it has been accomplished for some countries with sufficient accuracy. We can also measure the general range of prices. We know, therefore, whether a given sum of money incomes at one time means more than a given sum at another time. If we know that money incomes have increased, and that the range of prices is unchanged, we are sure that real income, in terms of consumable commodities, has increased.

Further, we can do something toward measuring "real" income directly. We can ascertain what has been the consumption per head of population, at different times, of such articles as flour, sugar, tea, coffee, cotton, wool, and the like. The results give significant indications regarding the increase of income in terms of commodities. We know that the average consumption of such things has much increased in recent times and that material welfare has so far advanced.

But how far total utility or "psychic income" has increased we have no accurate notion. We may feel sure that it has increased in some degree; but whether in the same degree as consumer's wealth, or in less degree, or even in greater, we do not know. We cannot measure how great total utility was before the increased supply of economic goods or how great after. The supply of the things

which minister to enjoyment can be measured but not enjoyment itself. Virtually all problems of legislation and applied economics can be settled, and habitually are settled, according to the results in terms of the former sort of income. Hence we do best, for almost all economic reasoning, not to go beyond the tangible and measurable facts of consumer's wealth. Even tho consumer's goods be but a sort of capital, and even tho total utility be in the last analysis the true income, the only kind of income about which we can reach results of quantitative accuracy is that "real" income which consists of enjoyable things.

§ 7. The principle of diminishing utility, if applied unflinchingly, leads to the conclusion that inequality of incomes brings a less sum of human well-being than equality of incomes, and that the greater the inequality, the less the approach to the maximum. If additional increments of any commodity yield less enjoyment than preceding increments, the same is true of increments of income in general. A man who already has five oranges gains less from a sixth than he who has but one orange gains from a second. A man who has an income of $10,000 gains less from an additional $100 than does the man who has an income of $1000. This is stated in another way in the proposition that gambling between persons of equal income always brings an economic loss. If two men, each having $1000, bet $100, the gain to the winner from the increase of his possessions to $1100 is less than the loss to the loser from the drop of his possessions to $900. All this follows directly from the hedonistic calculus—from the principle of diminishing utility.

It hardly needs to be said that this hedonistic calculus does not tell the whole story of human happiness. One of the unfailing sources of satisfaction, deep-rooted in human nature, is the response to the instincts of emulation and distinction. But distinction implies inequality. Tho there may be distinction and inequality in other ways—in rank or fame—difference in economic possessions has been an immense stimulus and an immense resource to almost all men. Much of the spice and flavor of life would be gone with flat equality.

None the less, it remains true that there is an opposition between inequality and maximum happiness. The opposition be-

comes obvious when the inequality is very great. High disparity of incomes means a net loss in enjoyments; the rich gain less than the poor lose. Tho some emulation and distinction be essential to a full and happy life and tho some inequality of income be a persisting consequence of distinction, such great inequalities as are familiar in modern society, and indeed in all societies advanced much beyond barbarism, cannot possibly bring the most effective distribution of the material sources of enjoyment. The consciousness, more or less obscure, of the inconsistency between maximum well-being and great inequality underlies the whole modern social movement; this movement has for its main goal a more equal distribution of income. From this flow the characteristic tendencies of our time—curbing of monopolies, extension of government industry, labor legislation, progressive taxation; last but not least, socialism. Inequality may be and probably is an indispensable spur to the full application of men's best faculties, and an inevitable outcome of free and vigorous industry. But *prima facie* it does not lead to the best distribution of well-being. It is always on the defensive; and the greater and more persistent it is, the more difficult is its defense.

§ 8. The conceptions of marginal utility and total utility do not explain anything and everything in the exchange of goods. They are not to be made a fetish. They bear most on the play of the primal wants for food, clothing, shelter, for simple substantial comforts. As regards the wide and growing range of things which cater to the more sophisticated conventions and demands of men, and especially to that love of distinction which apparently is inborn and ineradicable and certainly plays queer tricks, the conception of total utility seems to be most significant within a moderate range of supply, and seems not at all to explain what happens when the supply is much greater and time longer. Measurement of utility or welfare is least of all possible for what we deem the higher things of life, literature and music, the drama, painting and sculpture, architecture; not to mention religion and the altruistic spirit. As the industrial arts advance and the ordinary means of satisfaction become more abundant, as more and more of men's labor and contrivance are turned to other than the

primal wants, the play of demand becomes irregular and even capricious; and this holds even for the wants of the lower range. At the same time, as the sense of common dependence and common interest becomes greater, individual wants and satisfactions tend to become less important than those which are felt by all to be for the good of all. What promotes well-being and happiness comes to be in less degree a matter of the maximum satisfaction of the individual and in greater degree one of the happiness of groups, communities, nations.

The economist may put such reflections aside, not because he thinks them irrelevant or unimportant but because they do not come within his range. Yet they cannot fail to obtrude themselves on him and on his readers. Relevant they are, and important; set aside merely because calling for a different sort of approach and perhaps a different type of mind from that which aims primarily at some exactness in the quantitative measurement of the consequences of man's work and wealth.

CHAPTER 10

MARKET VALUE. DEMAND AND SUPPLY

§ 1. The conditions of demand and the demand curve.—§ 2. Demand possibly discontinuous, usually continuous. Elastic and inelastic demand.—§ 3. How value is determined by marginal vendibility, for a fixed supply. The equation of demand and supply.—§ 4. A varying supply: the equilibrium of demand and supply.—§ 5. How far the supposition of a fixed supply, how far that of a varying supply, conforms to the facts. The circumstances that act on daily and on seasonal prices.—§ 6. Qualifications as to the market value of capital goods.—§ 7. Retail prices seem to follow wholesale prices, but in the end govern wholesale prices. The advantage of fixed retail prices.—§ 8. Current market prices are what people commonly mean when they speak of "fair" prices.—§ 9. Sporadic cases where value is affected by utility to sellers.

§ 1. In the preceding chapter the first principle of value has been stated. The value of an article depends on its marginal vendibility. It is the price at which the last installment can be disposed of—the price that settles, in turn, under the ordinary conditions of competition in the market, the price at which the whole supply will be sold. It remains to illustrate this principle further, and to explain in what manner it operates in the complexities of actual life.

Let us first illustrate the main principle graphically. On Figure 1, prices are measured along the perpendicular axis OY; quantities, i.e. the several installments offered in the market, are measured on the horizontal axis OX. Let it be supposed that the first dose, say of sugar, is represented by the horizontal line OA, and that this dose would command the price OP. Its value would then be indicated by the area $OPA'A$—the quantity multiplied by the price. Suppose now a second dose to be offered, indicated by the line AB. Under the influence of the principle of diminishing utility its price would sink to OP', and the whole supply would now be sold at this price (or rather, as will presently be explained, at no higher price than this). The total value of the increased supply would now be indicated by the area $OP'B'B$. Add now another dose, the supply being OC; the price sinks

again, and the value of the whole supply is $OP'''C'C$. And so on, with the supply OD, the price will be OP''', and the whole value $OP'''D'D$, and with the supply OE, the price will be OP'''' and the whole value $OP''''E'E$.

Strictly speaking, under the conditions here assumed, we should not know that the price for the quantity OB, for example, was

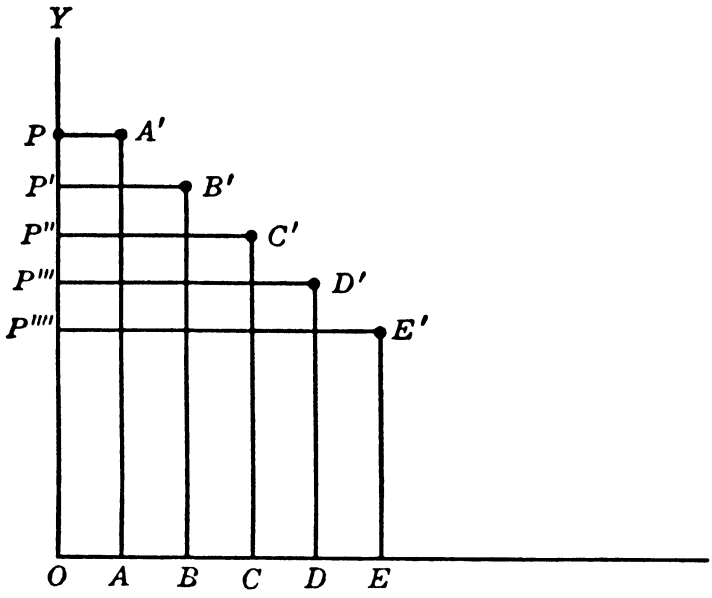

FIG. 1

fixed at the amount indicated by the lines OP' or BB'. We should only know that it was not higher than OP' and not lower than OP''' (CC'). In order to induce the supply OB to be taken off, the price must be at least as low as OP''; otherwise the buyer would not take it. But if the buyer offered less than OP'', the seller would still rather dispose of his supply than have it left on his hands; and until another potential buyer came on the scene, there is no telling what price the seller might not accept. But if another buyer comes, to whom the dose has the utility measured by OP''', and who is willing to pay the price so measured, the seller can compel the second buyer, stationed at B, to pay at least as much as the third competitor, stationed at C, would offer. Price, therefore, would be somewhere between OP' and OP'', or somewhere between BB' and CC'. So in each of the successive stages. The price must be at least low enough to

tempt the last buyer who must be called in to dispose of the whole supply offered. It may go a bit lower than this, until the point is reached at which a new buyer would enter and prevent the more desirous buyer—the more "capable" buyer, as he has sometimes been called—from beating the seller down. If there be a considerable difference between the utilities of the installments to successive buyers, there is a considerable range within which price is indeterminate.

§ 2. In the ordinary course of business dealings there are no such abrupt stages in demand as the preceding diagram (Fig. 1) assumes. There are not a paltry half-dozen purchasers, and a few pieces on sale, for any given article. There are many buyers, to whom great supplies are offered. Among the many buyers there are always some just ready to step forward, to whom the utility of the additional dose is only a shade less than was the utility of the previous dose and who are therefore called into the active purchasing market by the lower price. This situation is described by saying that demand is continuous. Where there are gaps between the utilities to different purchasers, and consequently between the prices they are willing to pay, demand is discontinuous. For most dealings in modern communities the points are so near together—the gradation of utility and demand is so close—that they may be represented as joined into a line or curve. That curve has a smooth downward inclination from left to right, like the unbroken line *DD'* in Figure 2. It indicates that successive doses of any article have gradually diminishing vendibility and must be offered at prices that insensibly become lower and lower as greater quantities are disposed of. It is called the *demand curve.*

The curve indicates the nature of the demand for a commodity. The *rate* at which the quantity purchased—for brevity it may be called the consumption—responds to changes in price will differ according to the commodity, and for any particular commodity the responsiveness of consumption to changes in price may vary over different sections of the curve. If the demand curve descends slowly from a particular point—as does the dashed line *ee'* from the point *P* in Figure 2—it indicates that with a small decrease in price consumption increases rapidly. If a small change in price

leads to a change in consumption greater in proportion than the change in price, i.e. if demand is very responsive to a change in price, as is indicated by the dashed line ee' in Figure 2, the demand for the commodity is said to be elastic. On the other hand, a curve descending more quickly from a given point, as does the broken

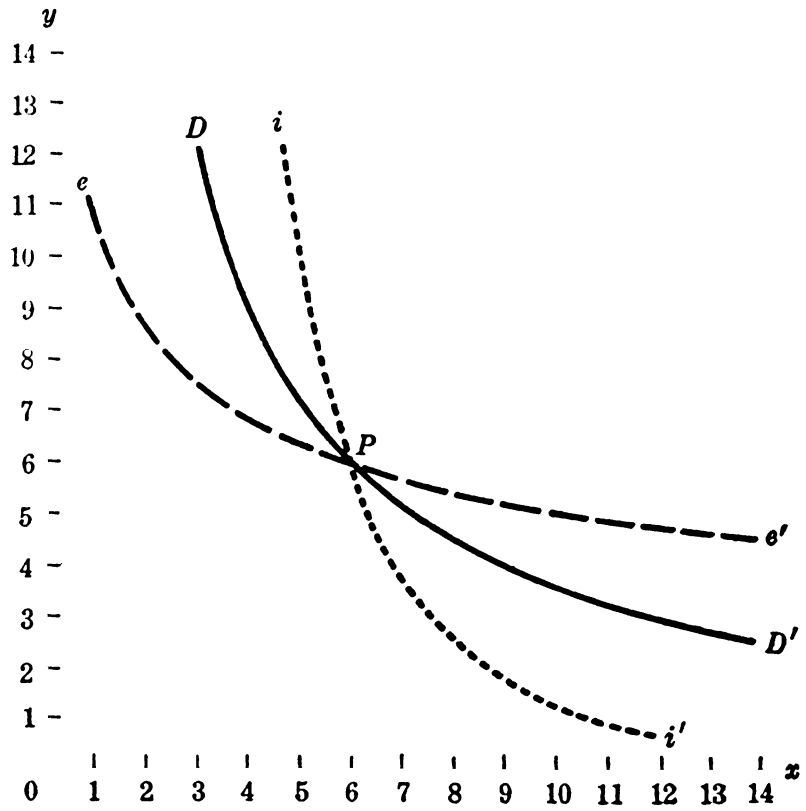

FIG. 2

line ii' from the point P in Figure 2, indicates that utility or purchasing power diminishes rapidly; additional purchases are not made readily with a decline in price, and the decline in price with a given increase in supply is more abrupt than in the preceding case. In this case the demand for the commodity is said to be inelastic; consumption does not respond readily to a lowering of price; the proportionate increase in consumption, with a small change in price, is less than the proportionate decrease in price. The cause of inelasticity must be in some degree a rapid diminution of the utility of added installments, but the cause may be in large part an inequality in means. If some purchasers are very rich, others

well-to-do, many others poor, commodities may meet a highly inelastic demand in the market but not necessarily suffer a corresponding diminution in their power of yielding enjoyments to mankind.

If we were to use the term "inelastic" in strict accord with its ordinary connotations, we should say that demand was inelastic only when the quantity purchased remained the same whatever the price. The demand curve would then be a perpendicular line. And we should say that demand was elastic if the quantity purchased increased even to the slightest extent as price declined. Since every commodity (the exceptions are negligible) is bought in somewhat greater quantity at a lower price, the demand for every commodity would then be elastic. In this meaning of the term there are degrees of elasticity, but inelasticity of demand hardly ever appears. The common usage of writers on economics, however, has given a different meaning to "elastic" and "inelastic" demand. In this usage the curve *DD'* on Figure 2 is a curve of unit elasticity, while the curves *ee'* and *ii'* are curves respectively of elastic and inelastic demand. The dividing line between elasticity and inelasticity is placed at the stage where the elasticity of demand is unity. Imagine a commodity for which the same identical sum is always spent by purchasers. The quantity purchased does indeed increase as price falls, but increases in exactly such proportion that the quantity multiplied by the price always yields the same product; and, on the other hand, while the quantity purchased becomes less as the price rises, the diminution is such that at the higher price per unit the total spent still remains unaltered. This sort of case we describe by saying that the elasticity of demand is unity. Compare such a commodity with one of which the quantity purchased increases greatly as price declines—so greatly that the total spent at each several stage is greater than at the preceding stage. Elasticity of demand is then greater than unity. Conversely, if there be still another commodity of which the quantity purchased, tho it increases as price declines, increases so slightly that the total amount spent at each several lowered price becomes actually less than at the preceding stage, then elasticity of demand is less than unity. It is convenient

to give precision to the difference of degree by speaking of the first sort of case as showing an elastic demand, and by speaking of the second as showing an inelastic demand.[1]

The demand for necessaries is inelastic. Nearly the same quantity of bread will be bought, whatever the price. No doubt a high price will in some degree check consumption, and a low price will lead to more liberal and careless use. But when the indispensable supply has once been got the decline in utility because of greater quantities is rapid. For articles of this sort a comparatively small shortage in supply will cause a proportionately large increase in price, while a comparatively small redundancy will cause a proportionately large decline. The curve *ii'* is the graphic representation of the inelastic demand for necessaries and of the relatively large fluctuations in price induced by relatively slight changes in supply.

Any article which, tho not necessary, is yet clung to with persistence by consumers has similarly inelastic demand. Meat for example, tho not a necessary, has an inelastic demand among the well-to-do. On the other hand, the substantial comforts of life—things not indispensable yet prized by all the world—often have an elastic demand. Such are those articles of food which, tho not necessaries, please by their flavor and variety—sugar, fruits, vegetables, tea, coffee. For almost all except the well-to-do meat is such an article. In the upper part of the supply it has an inelastic demand, in the lower part a very elastic demand.

In general, elasticity of demand is increased by an equal distribution of wealth, while an unequal distribution leads to inelasticity in demand. This effect of inequality illustrates once again the caution which needs to be observed in applying the principle of diminishing utility to the phenomena of value as they appear in modern communities. If all people had the same incomes diminishing utility would be the one cause acting on the elasticity

[1] The demand curve of a commodity for which the elasticity of demand is unity, is a rectangular hyperbola. It is a curve such that every rectangle drawn parallel to the axes (asymptotes) of our figure, of which the corner impinges on the curve, has the same area. This is the nature of curve *DD'* in Figure 2, in which curve *ee'* indicates elasticity of demand greater than unity and curve *ii'* indicates inelasticity.

of demand, and the inclination of the demand curve would be significant of the rate of diminution in the enjoyments yielded by successive increments. In fact, the demand curve is much affected by the circumstance that persons of means can afford to bid high for the first increments, while the great number of those with small means cannot bid until a low price is reached. The lower bids of the latter mean not so much a diminution in enjoyments as a scantness of money means.

§ 3. We proceed to consider how the mode in which the value or price of an article is determined at any particular time—the problem of market value.

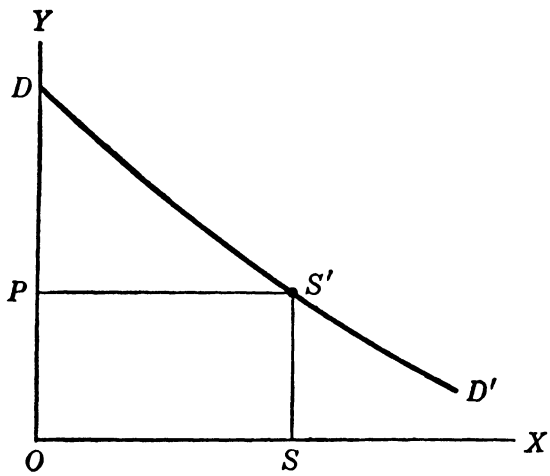

FIG. 3

Suppose the supply of a commodity to be fixed; suppose it to be offered on the market by competing sellers; suppose it all to be offered without reserve. Then the value of that commodity will be determined by its marginal vendibility. If all is not sold at that price by the competing sellers, some part of the stock will not be disposed of. This situation is graphically represented in Figure 3. Given an amount *OS*, the resulting price will be at the point where the perpendicular line *SS'* will cut the demand curve *DD'*. That line (*SS'* = *OP*) measures the marginal vendibility of the supply *OS*, and so measures the price at which that supply will be sold.

The total exchange value of the supply is indicated by the area *OPS'S*—the supply multiplied by the price. Total utility is indicated by the irregular area *DOSS'*; consumer's surplus by the (more or less triangular) area *DPS'*. Those purchasers who, rather than go without the article, would have been willing under stress to pay a higher price than *SS'*—as high as *OD*—secure some surplus of satisfaction.

The same proposition regarding the mode in which the value of

an article at any given time is determined was stated by the older writers in a somewhat different way. They said that market value was settled by the equation of supply and demand. The everyday way of putting it is to say simply that the value of a thing is determined by supply and demand. This is loose, since it implies that supply and demand are causes that act independently, and are not themselves influenced by price. Demand in the sort of case here supposed is certainly affected by price. The lower the price of an article, the more of it will be demanded; the higher the price, the less will be demanded. To say that price depends on demand, therefore, seems to be reasoning in a circle, since, if price is affected by demand, demand is no less affected by price. Hence the more careful phrase just quoted: the equation of supply and demand. Given a *fixed* supply, there is one price at which the quantity demanded will be just equal to the fixed quantity supplied. To assume that there is one such price, and not more than one, is to assume continuity of demand, as explained in the preceding section—an assumption that holds good of the vast majority of articles bought and sold in the markets. This one price evidently represents the marginal vendibility of the supply. Tho the phrases "marginal utility" or "marginal vendibility" were not used by the older writers, their statement in terms of an equation of demand and supply is substantially the same proposition as the more modern one which reasons on the basis of diminishing utility, marginal utility, and marginal vendibility.

§ 4. In both of these statements of the principle of market value—the older one of an equation and the newer one of the marginal vendibility of supply—the underlying assumption is that a fixed quantity is put on the market. But is this assumption tenable? Does it conform to the usual state of facts? We have just said that demand, in the sense of quantity demanded, is not independent of price. Is not the same true of supply? In the ordinary case it is hardly accurate to say that the quantity offered in the market is fixed, and is independent of price. As price goes higher more sellers will be tempted to offer their wares, and supply will become larger. As prices go lower, supply will become

smaller. Must not the theory of market value be adjusted to variable supply as well as to variable demand?

In some instances the supposition of a fixed supply is clearly in accord with the facts. When a large crop of strawberries comes on the market it must be disposed of once for all. There is no keeping back any part of the supply of a perishable commodity. The total quantity on hand must be disposed of for what it will fetch—for the marginal price. Not very long ago the list of commodities of this kind was a large one; it included fresh fish, all vegetables and fruits, even meat. Modern improvements for the preservation of nearly all such things, thru cold storage and canning, have greatly shortened the list. Most commodities are not put on sale with headlong suddenness. They are offered in installments. They come into the market in a flow or stream, not as an abruptly offered stock. The rate at which they come in and the amount which will be offered at any given time depend on the price. A higher price quickens the flow and leads to larger supply; a lower price checks the flow.

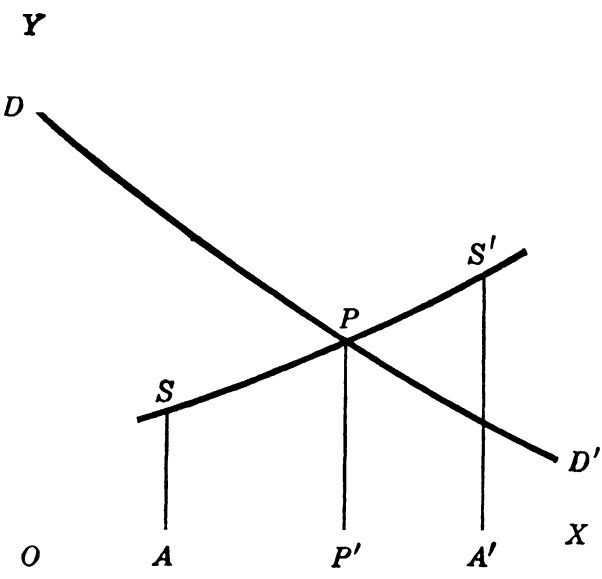

FIG. 4

It is not difficult to adjust the theory of market value to the case of variable supply. On Figure 4, let *SS'* represent the conditions of a supply that varies with price, becoming greater as price rises and smaller as price falls. Here, as on the previous figures, quantities are measured horizontally along the axis *OX* or parallel to it, and prices perpendicularly along the axis *OY* or parallel to it. At the price *SA*, we may suppose the quantity *OA* to be forthcoming on the market. As the price rises the quantity increases. At the price *PP'*, the quantity offered is *OP'*; at the price

S'A', the quantity offered is *OA'*. Evidently the line *SPS'*, which is the supply curve, has an upward inclination, the reverse of the inclination of the demand curve *DD'*. A rise in price, which causes the quantity demanded to become less, causes the quantity offered to become greater.

The supply and demand curves, moving in opposite directions, must meet, and in our figure they meet at *P*. The price *PP'* is the equilibrium price, the market price fixed by the play of varying supply and demand. At that point the quantity offered is equal to the quantity demanded: the equation is satisfied. If a higher price is asked, the quantity demanded will be less and the quantity offered will be greater. Sellers will put on the market more than buyers will take; price will fall; some sellers will then withdraw and some buyers will come in, until equilibrium is reached. And so in the reverse case: at any lower price some sellers will withdraw, some buyers will be tempted in, and readjustment will again bring the price to the point of equilibrium *PP'*.

§ 5. It has just been said that of these two modes of statement—the one proceeding on the supposition of a fixed supply, the other on that of a variable supply—the second is more in accord with the facts. Yet the first likewise is so in accord. Both must be had in mind for an understanding of the course of prices in a market.

On any given day, in a well-organized market, the actual settlement of market price undoubtedly takes place thru an adjustment of supply as well as thru a response from demand. On the cotton exchange or the produce exchange, or in any place where brokers and dealers meet, a process of higgling and bargaining goes on. More or less of the article is offered and demanded, with fluctuations in prices which are usually within narrow limits on any one day and which result in an equilibrium price for that day. But this daily equilibrium price is itself affected by an underlying and more important equilibrium price. While the amount which is offered in the market from day to day—the supply—varies considerably, and varies in response to changes in prices, the total amount which can be supplied over a large period usually is fixed. Take, as a typical case, the price of cotton, which fluctuates on the exchanges from day to day in response to the ever-changing

play of offer and demand. The total amount of cotton available for the season is not a variable quantity. It is so much and no more, depending on the crop of that season. The price at which the whole will be disposed of depends on its marginal vendibility or on the equation of supply and demand (whichever mode of statement be preferred) and is the outcome of a total supply which is fixed. The fluctuations in price from day to day oscillate about this seasonal equilibrium price.

Still using the cotton market and cotton prices for examples, we may note that, while the supply for the season is fixed, no one knows in advance with certainty just how great that supply is; still less at what price the supply, even if accurately known, would be disposed of. Hence a period of uncertainty, of rumors and guesses, of selling and buying by brokers and dealers and manufacturers, by any one who chooses to operate on the cotton market—in short, all the phenomena of speculation. Cotton in the United States (the crop in this country dominates the world market) is picked in the autumn, and the amount harvested is known by December 1. But thruout the summer months there are reports of the condition of the growing plants, which foreshadow, tho with uncertainty, the amount of the coming crop. During the picking season more and more certainty is reached. Finally, under modern methods of gathering such information, the amount comes to be known with great accuracy. Then rises the question to what degree the price will be affected by the amount. It is certain that a small crop will command a higher price, a large crop a smaller price. But the conditions of demand or consumption are fluctuating from year to year, no less than the supply from the crops. Just what will be the seasonal equilibrium price for a crop of a given size, no one can say in advance. It is reached by a succession of tentative market prices. From day to day, and from month to month, the market price is settled by the adjustment of variable amounts offered in the market by dealers. For the season, it is settled by the adjustment of a fixed supply to the marginal price at which the whole will be disposed of.[1]

[1] It need hardly be said that "normal" conditions of marketing are here assumed, under which the whole of a cotton crop is sold shortly after it is ready for the

It is not to be supposed that even on a single day is there one price rigidly settled by the equilibrium of demand and supply. Even in the most highly organized markets there may be simultaneous sales at different prices, and where there are newly discovered conditions affecting the seasonal range, such as a crop report, there may be considerable fluctuations in the course of a day. These oscillations give the opportunity to the astute bargainer. Some buyers, not cool-headed enough to bide their time, will pay more than the equilibrium price. On the other hand, some sellers, unduly anxious lest their supplies be left on their hands, will sell at less. The shrewd and unexcitable person, carefully watching the course of dealings, may buy at one price from the over-eager sellers and sell on the same day at a profit to over-eager buyers. It is sometimes said that all the capital a speculator needs is a pencil and a block of paper and all the knowledge he needs is a knowledge of human nature. This is by no means the whole story; yet it is true that a certain faculty of judging human nature and an impassive demeanor are important in the equipment of the professional dealer, and play no small part in those speculative operations which are discussed in the next chapter.

The more the actual dealings in a market are confined to persons who are shrewd and well-informed the more probable is it that there will be an exact equilibrium price. And in any market where dealings are habitually conducted on a considerable scale there will be an equilibrium price which, tho not rigid, is maintained between comparatively narrow limits, and that price will represent the judgment then currently held of the probable seasonal price. Here, as in all economic analysis, we have to do not with hard and fast phenomena but with the wavering doings of human beings. For the sake of bringing out clearly the underlying general probability—a probability which often is so great as to be virtually a certainty—we state our reasoning and conclu-

market—picked and ginned. The conditions of the period 1933–39 have been quite different, and beyond the ordinary range of economic happenings: the withholding from the market and the storage of a great part of the domestic cotton supply by the federal government, which guarantees a price to the producers for the supply withheld. To such a situation the analysis in the text cannot apply. I undertake no guess as to the probable continuation of this sort of public action.

sions in semi-mathematical form, as in the diagrams and figures that have preceded. But it must be remembered that the conclusions hold good not with mathematical certainty but simply as statements of tendencies to which the actual market conditions more or less conform.

What is true of cotton holds of other agricultural commodities, whose supply also is settled by the crops of each season: of wheat, corn and other grains, of hay, flax and hemp, hops, sugar, tea, coffee. There is always a seasonal price, around which fluctuate the market prices for shorter periods. Something like this holds of other commodities also. It is true that agricultural commodities show more than most others the temporary fixation of supply. The supply of manufactured commodities changes more smoothly and continuously. The amounts offered in the market can often be increased and diminished without waiting for nature's process of growth. But even here there are important limitations. For any given period of moderate length—a half year or a year—there is something like a fixed supply. Iron, for example, is continuously produced, and the amount of production responds in some degree to the fluctuations in price. But the quantity available for any given period depends on the mines of iron ore and of coal which are open, and still more on the furnaces and works which are ready to smelt and shape the iron. The supply can be increased or decreased only with considerable difficulty. It will not readily decrease: continuous operation is a condition of almost any profit at all and the existing iron mines and works will be kept going unless the prospects for profit are very bad indeed. Nor can it be rapidly increased. New mines and works can indeed be added, but this takes time. Again, tho the output from the existing concerns does not come on the market at any fixed or regular rate, the whole of it is almost sure to be offered for sale within the current season of operations. Thus a seasonal equilibrium of supply and demand establishes itself. Around this seasonal price the current market prices fluctuate as varying amounts are offered and demanded from day to day and from week to week.

Sometimes dealers, looking far ahead, carry stocks over a con-

siderable period. In this way the supply on hand, even the seasonal supply, may be sensibly affected and the seasonal market price may be affected correspondingly. If for example the wheat crop in any year is very large, and the price unusually low, some dealers may withdraw considerable amounts from sale, store them, and plan to sell them at a profit in the next year, when a smaller supply and higher prices may be expected. But this is a risky operation. It involves the locking up of large money means. The next season may again bring a large crop. There is the possibility that the wheat held in storage may spoil and become valueless. The more durable a commodity the more possible is it that considerable parts of the supply will be carried over from season to season. If iron and copper are unusually cheap, stocks of them can be bought and put aside with a minimum expense for storage and with no risk of deterioration, and held in expectation of higher prices after a year or two. Yet even for these durable articles such operations seem to be uncommon. Most persons in active business, and especially dealers and middlemen, do not try to look far ahead. They study the conditions of the present and the immediate future and govern themselves accordingly. The withdrawal of stocks from the seasonal market seems to be no considerable factor in the play of demand and supply.

§ 6. Strictly speaking, the discussion of utility, marginal utility and of marginal vendibility applies to consumer's wealth only. Capital yields no utilities directly. Materials, implements, machinery are but means toward providing utilities at a later date. Their utility is a derived one, depending on the utility of the consumable goods they aid in making. Tho the principle of marginal vendibility works out its results for capital goods also, it does so thru an intricate process and with some complications.

For example, when the cotton crop is small, the price of cotton rises; marginal vendibility is greater, we say, for the smaller supply. But the cotton is sold by the planters and farmers first to the dealers; they sell to the manufacturers; these again, thru another set of dealers, sell the cotton cloth to those who wear it. It is the satisfactions got by these ultimate consumers that in the end determine the value of cotton for a given supply. But it is the

manufacturers who are commonly spoken of, in the language of the market, as the "consumers" of cotton. These are sometimes in a position in which they *must* buy cotton. They have a plant which must be run if it is to earn anything at all, and a force of workmen which, to remain efficient, must be kept together. Each manufacturer wishes to keep his plant working at full capacity and his workmen fully employed, yet with a small crop there is less cotton to be worked up. On the other hand, the extent to which the eventual users and buyers will pay at a higher rate for the diminished amount of cotton cloth is an uncertain factor. The manufacturers try to get from the merchants to whom they sell a higher price for cloth corresponding to the higher price of cotton. Both sets of business men will say that it is the high price of cotton which *causes* the high price of cloth. Yet the reverse is at bottom the case; only because the cloth can be sold at a high price does the raw material command a high price. How close the correspondence in price will be, how much the investments and commitments of the manufacturers will affect the situation, how the calculations and transactions of cotton dealers and speculators and cloth merchants and cloth buyers will act on prices at any one date and thru the season—these are matters on which the action of the fundamental economic forces is slow and uncertain. There are analogous complications when there is a very abundant cotton crop. Then manufacturers are not prepared to work up an unusual supply of the raw material; merchants and retailers are not certain how far and at what prices they can find a market for additional quantities of cloth. Tho cotton cloth is a commodity having an elastic demand, raw cotton, despite the fact that demand for it is derived from that for cloth, may show from season to season fluctuations such as one would expect in a commodity for which the demand is inelastic.

Other kinds of capital goods are to be used for durable tools and plant. Such are iron, copper, timber, brick, stone. In the end, the demand for these also rests on the utility of the enjoyable commodities made with them; they also have a derived utility. But proximately the demand for them is from persons who wish to use them in connection with new investments. When

the prospect of profit is good, the prices of these things rise; when the prospects are bad, their prices fall. Hence their prices are closely connected with those alternations of activity and depression, of good times and bad times, which are among the most puzzling of economic phenomena. It is true that their market price is settled by the amount which the last purchaser—the least eager of the buyers—is willing to pay. And in the end, no doubt, what that purchaser is willing to pay depends on what can be got in turn for the consumable goods made with the aid of the capital goods. But the chain of connection is a very long and irregular one, and the market price of the iron or copper is affected first of all by current expectations as to investment activity. It would be absurd to apply to these articles any strict principle of marginal vendibility. That principle, like others in economics, works out its results with all sorts of qualifications and complications.

§ 7. Retail prices might be expected to illustrate most clearly the play of marginal vendibility, for here enjoyable goods are sold to their consumers and the utilities from them are nearest realization. Yet in fact retail prices appear to be less subject to the working of supply and demand than wholesale prices.

Retail prices are governed proximately by custom. People pay the traditional or going price. Even the amounts which they purchase appear to be governed by custom; they buy the quantities which they are in the habit of consuming. And the retail prices which establish themselves as customary seem to be governed by wholesale prices. The retail dealers charge more when there is a considerable and apparently definitive rise in wholesale prices; and competition among themselves causes them to charge less when there is a considerable and lasting fall. No doubt, the accommodation of retail to wholesale prices may be slow. When wholesale prices rise, shopkeepers hesitate to ask more, partly because each one fears that his rival may entice a customer away by keeping to the old price for a while. Conversely when wholesale prices fall, no shopkeeper willingly gives his customer the benefit of the change; he waits until some competitor precipitates it. The two sets of prices in the end move together but changes in retail prices seem to follow rather than precede those in wholesale prices.

But all this is in appearance only. The consumption of every commodity is affected by its price. A rise in price checks purchases, a fall in price stimulates them. Tho it would appear that people continue to buy simply what they are used to buying, this is true only of buyers who are above the margin—those who have been enjoying a consumer's surplus. There are always others just on the margin, to whom at the ruling price the purchase is worth while and who cease buying when the price goes up. And conversely, when price falls there are always some additional purchases. How great the changes in consumption are with rising or falling price, depends on the elasticity of demand, and some degree of sensitiveness there always is. So certain is this that the wholesale dealers reckon on it in advance, and at once accommodate the current prices in the wholesale market. It is they who usually are best informed regarding the general situation. They know when a crop is short, or a new source of supply has been opened, or an invention is cheapening production and increasing the amount offered in the market. It is they, too, who can best observe when the habits of consumers are undergoing change and so are affecting the purchases of a commodity. In case of an increase in demand, any one retailer may indeed notice that his customers are buying more than before. But this may seem to him an isolated phenomenon; he simply orders more from his wholesale agent and expects to sell more at the old price. When, however, orders from many retail dealers thus come in to many wholesalers the market responds and price goes up. The retail dealer then charges more to his customers because he has paid the wholesaler more for his goods, the real influence at work being that the customers, taken as a whole, want the goods more. Here, as in all the phenomena of value and price, the stocks held by dealers, whether retail or wholesale, have an effect in preventing abrupt changes, and obscure and delay the restoration of the equilibrium of supply and demand. In the end, that equilibrium, resting on the demand of the marginal purchaser and so on the principle of marginal vendibility, settles both wholesale and retail prices.

In the earlier stages of industrial life and even in many countries which have attained a comparatively advanced stage, retail prices

are fixed by a direct process of higgling between sellers and buyers. In the very earliest and most primitive stages, exchanges are few and sporadic and higgling plays a very important part. There is then nothing in the nature of a market price or customary price; and the astuteness of the bargainers, the needs and whims of the moment, even the possibility of physical force, affect the terms of exchange. As the division of labor is extended farther, and continuous exchange and sale develop, something like a market price establishes itself. That market price is likely soon to become a customary price, representing roughly an equilibrium of current demand and supply; yet, tho customary, it is likely also to be subject to bargaining, and to vary more or less from the customary rate.

In the highly developed countries of modern times, bargaining in retail dealings has been almost entirely discarded. The dealer sets a price at which he will sell and at that price the purchaser may take the article or leave it. The tacit understanding is that the price so fixed shall be the current or market price, and that it shall be the same for all customers at the shop. The practice of fixed prices saves a vast amount of time and friction. The purchaser need not be on the watch to discover what other dealers are asking and what is the going price; while, if he is not a marginal purchaser, but is enjoying some consumer's surplus, he need not be on his guard lest the dealer take advantage of his potential demand. The ease of everyday purchases and the efficiency of labor in retail operations are immensely promoted. Retailing on a large scale, conducive as it is to economy of labor, would be impossible without the practice of fixed prices. In many parts of the continent of Europe it has not been fully adopted. There the retail dealer still asks not the price which he will take once for all but a price which he hopes to get from the individual purchaser, and which he is prepared to lower if the purchaser bargains shrewdly. The result is friction, waste of time and inefficiency.

§ 8. The current market rate is what people usually have in mind when they speak of a "fair" price. This is what the retail dealer is expected to charge as his fixed sum. If he asks a higher price than is usually asked at the time by other dealers for the same

thing—still more, if he asks a higher price from one purchaser than from another—he is said to be charging unreasonably, or overreaching, or even cheating, and he is likely to lose his custom. There is often a similar attitude in regard to wholesale prices. Many large dealings in the wholesale market are concluded, in the great civilized communities, on the principle of fixed prices. A manufacturer or merchant in search of a given article orders what he wants from an agent or correspondent of established reputation, with the understanding that a fair price—that is, the ruling market price—will be charged. Here, as in retail dealings, confidence in honesty and acceptance of prices as they stand conduce to the easy dispatch of business. Underlying all, however, is bargaining somewhere, a more or less overt adjustment of price to supply and demand. What is a really just price at which goods shall be sold is a question much more difficult than is supposed by most persons who use the phrase. Indeed, few who talk of fair and unfair prices are conscious of the very nature of the problems involved. They are problems not of exchange but of distribution, taken up at a later stage.

§ 9. The discussion thruout the preceding pages has gone on the assumption that utility to the buyer is the only aspect of utility that needs consideration. The seller is supposed to put his wares on the market once for all, and to dispose of them, sooner or later, on such terms as their utility to buyers makes possible. But may not utility *to sellers* also affect price, by affecting supply? May not part of the supply be withdrawn by the sellers for their own use? Would not the extent of this withdrawal depend on the price, and so introduce a further complication in the theory of market value?

It is entirely conceivable that utility to sellers should thus affect price. In the case of the five oranges, supposed above, it is conceivable that the holder of them might consider the possibility of enjoying one himself, and would be led to do so more and more as the price descended. At fifty cents he would readily part with one of his oranges, but at five cents he might conclude to eat one, and so withdraw part of the supply. And if we suppose not one seller with a few oranges but many sellers with many oranges,

and suppose that among these sellers there is a considerable possibility of withdrawals for consumption, we have a problem more complicated than that of sales based on utility to buyers only. A great deal of intellectual ability has been given by economic writers to the analysis of this problem and to the careful statement of the terms of exchange that would result under various hypothetical conditions.

But almost all this subtle analysis is in the air. Under a developed division of labor, utility to sellers does not affect value. Men produce with no reference to their own consumption. They produce for the market. The supplies in their hands of the things made by them are so great that the importance to them of any unit is *nil*. They throw their product on the market without reserve. No doubt, if that product were very great indeed—such as to make the marginal utility to purchasers almost *nil*—the sellers might stop to consider whether they could not use some fraction of it themselves. Farmers may consume more apples when a very heavy crop causes apples (on the trees) to be nearly valueless. But any supply created by effort and with a view to sale is rarely so far increased that price sinks near zero, and where by mischance price is very greatly lowered the effect of utilization by the makers (sellers) is so slight as to be negligible. In the ordinary case virtually the whole supply is offered once for all on the market.

The case would be different if supplies got into people's hands from the start without reference to sale and disposal. If they were rained down from heaven in small amounts, prices would be affected by utility to sellers quite as much as by utility to buyers. We may imagine that in early times, before division of labor and exchange had developed far, sporadic exchanges took place under these apparently simple tho really complex conditions. But they must have taken place either with very vague consciousness of utility or under the influence of customs which greatly affected the actual terms of exchange. Ingenious hedonistic calculations probably throw little light on what happens in the stray exchanges of barbarians.

There are in the modern world occasional cases where exchange is affected by utility to sellers. When a fine old picture or a family

heirloom is put on the market its price may depend much on the attachment which the owner feels for it. Articles of this sort, of sporadic and limited supply, are in any case largely indeterminate in value, since buyers are few and demand is discontinuous. Their price may be made still more indeterminate by the fact that the seller (or sellers) may set store by the few specimens. The same is true, tho in very much less degree, of dwellings adapted to individual tastes. The ordinary house, planned like many others of its class, comes on the market under nearly the same conditions as other goods of homogeneous supply. But an odd house, built to suit the owner's idiosyncrasies of taste, stands more or less by itself. Its selling price may depend not only on the going price for houses of this range of desirability as estimated in the general market (that is, as estimated by buyers) but also on the attachment which the owner has for this particular one.

CHAPTER II

SPECULATION

§ 1. The fundamental effect of speculation by "legitimate" dealers is to mitigate fluctuations.—§ 2. Dealing in futures lessens price fluctuations of goods.—§ 3. Exchanges; standardizing.—§ 4. The evils of speculation, especially by "outsiders"; gambling; unproductive labor.—§ 5. Stock exchange speculation.

§ 1. THE phenomena of speculation connect themselves with the settlement of market prices. Something more may now be said on the good and ill of speculative dealings.

The term "speculation" is used in various senses. Often it implies the buying and selling of things by a person whose main business in life is different—"dabbling" in the market by "outsiders." But as often it implies buying and selling by persons who expect to make their living or their fortune by dealing in one commodity or in certain sets of commodities—persons who are "professional speculators." These again are sometimes distinguished from "legitimate" dealers, such as the wheat merchant or the cotton factor, who buy and sell a commodity year in and year out and are permanent middlemen for those who have it to sell and those who wish to buy. Between these various sorts of persons there are insensible gradations. All their operations have effect in determining market price and all are more or less in the nature of speculative dealings.

The fundamental effect of mercantile speculation—the kind last mentioned—is to promote the establishment of the equilibrium of supply and demand. It tends to make daily market prices conform to the seasonal market price, and to make the seasonal market price such that the whole seasonal supply is disposed of. Those who are skillful and painstaking in estimating the seasonal supply, and are shrewd and experienced in foreseeing the effect of a given supply on price, are the persons who are likely to make money in speculation. They buy when others offer at a price lower than the facts of the market warrant; they sell when others

bid a price higher than the facts warrant. The more the dealings of the market are confined to buying and selling between such shrewd and experienced dealers, the more likely is it that the seasonal price will be quickly and smoothly reached, and the less will be the fluctuations in price. With the inevitable uncertainties as to the amounts of the forthcoming supplies and the conditions of consumption and demand, there will always be differences of judgment between even the most expert dealers. There will be fluctuations in price, some ups and downs, some unexpected gains and losses—"speculative" profits or losses. But the general effect of speculation by such persons is to lessen fluctuations and promote the smooth course of exchange and consumption.

This lessening of fluctuations is advantageous alike to the ultimate consumers and to those manufacturers who in business parlance are often spoken of as the "consumers" of a raw material. For the ultimate consumers, say of wheat, the early and exact adjustment of price brings more even utilization of the available supply. If the crop be short, some lessening of consumption is inevitable, and it is better that the deficit be spread thru the season. The sooner and the more exactly the higher price is reached, the more likely is this result. Conversely, a large crop is better sold at a low price thruout the season than at prices ranging from high to low as the season progresses.

The good effect of speculation in this direction has been illustrated from the experiences of older days, when wide fluctuations in the price of food were common. Under modern conditions, with great areas of supply brought into competition by railways and steamships, abrupt changes in the supply of most foodstuffs and raw materials are rare. A poor crop in one country or section is likely to be offset by a good crop elsewhere. The seasonal supplies do indeed change and prices go up and down under their influence, but the variations are seldom great. But under such conditions as existed under the limited geographical division of labor before the eighteenth century, great fluctuations were common. Then the area from which any district or city got its food and materials was strictly limited. A crop deficiency meant a short supply, and necessitated the adjustment of consumption to that short sup-

ply. The dealers or speculators or "forestallers" who secured the supply and at once demanded high prices for it brought about the inevitable adjustment and caused a more even utilization of the stock in hand. All this was reasoned out by some of the older writers on economics, and led them to a warm defense of speculators and to a condemnation of laws aimed against speculation. Very likely their defense of speculation was carried too far. The process of buying from the farmers did not necessarily take place under active competition by the dealers or speculators, nor did that of selling to the consumers, and the gains of the speculators were enhanced by the ignorance or heedlessness of both farmers and consumers and might easily be thought larger than could seem reasonable. We know very little of the details of what took place in these early days and are prone to project into them ideas or conclusions based on our own experiences. But none the less it is probable that even in those times the influence of speculation was in the main to lessen fluctuations and promote the expedient rate of consumption. It is certain that this is its tendency under the modern conditions of wide markets, full information, active competition.

The development of cold storage in recent times has led to precisely this sort of equalized distribution of supply, under the influence of dealings that are essentially speculative. Fruit, meat, fish, eggs, no longer come on the market in spasmodic and irregular amounts. Supplies that are heavy at one time are bought by dealers, put in storage, and held for sale at a later period of scantier supply. Prices are more equable and on the whole the profits of dealers are probably less. There is less risk to them and the community gets its supplies at a smaller charge for the middlemen's services.

§ 2. The process of lessening fluctuations and distributing risks is probably promoted by the practice of dealing in "futures"—a practice with which the term "speculation" is especially associated. Goods are bought and sold not only for immediate delivery but for future delivery as well. The dealer who undertakes to deliver in the future a certain quantity of wheat at a certain price may not have in his possession the goods he sells; indeed, in the common course of such dealings in the modern markets, he usually does not

have them. He gauges the probabilities of the future and undertakes delivery on the terms which those probabilities suggest. Virtually, he guarantees a certain price for the future and takes his chances whether the guarantee will bring him gain or loss. The buyer is then relieved of the risk. The advantage of this elimination of risk is easily seen. The miller, for example, may wish to close a contract for the sale of flour in the future. By securing the needed wheat at a guaranteed price he is freed from all the risk of ups and downs and can give his undivided attention to his proper business of manufacturing flour.[1]

Hence it has happened, since the establishment of exchanges and the development of their varied operations, that millers carry on their business with a much smaller margin of profit than formerly. The difference in price, weight for weight, between wheat and flour is much less than it was thirty or forty years ago and the public gains in so far. When, for example, the flour-milling industry was first established at Minneapolis—where the falls of the Mississippi supplied power for grinding the wheat of a region singularly adapted to its growth—the possibility of profit for the miller was great. But he then underwent also the chances of loss from fluctuation in the price of wheat. As the exchanges developed and with them the practice of dealing for future delivery, he was able to free himself from these chances. The consequent regularity and solidity of the industry contributed to its systematic development on a great scale and so to the cheapening of flour. Inventions and improvements no doubt contributed greatly, but the elimination of market risks had an important share in reducing the difference between the price of wheat and the price of flour. Both in merchandizing and in manufacturing the growth of large-scale transactions, tho it has increased the gains of those individuals

[1] Even if he is not contracting for the future sale of flour at a given price but is simply manufacturing continuously for the market, he can escape by this same mechanism from the risk of fluctuations in the price of wheat. When he buys a given quantity of wheat to be ground into flour, he can sell for future delivery the same quantity of wheat. Thereafter, as wheat goes up or down, he loses as much by the one of these transactions as he gains by the other. The fluctuations no longer trouble him. This is a common practice among "conservative" millers. Cotton manufacturers also are getting more and more into the practice of thus "hedging" in their purchases of raw cotton.

who have the ability to carry on large operations, has lessened the margin between buying price and selling price and so has operated to lower prices for the consuming public.

The dealer or speculator who has sold for future delivery does not usually run all the risks of the transaction himself. He is likely before long to buy from another dealer, for future delivery, some part of what he has contracted to deliver, perhaps the whole; that other dealer in turn shifts part of the business to a third, and on. The process of gauging the course of the market fluctuations is hardly ever carried thru the whole of a season by one person for any one transaction. The dealers constantly buy and sell among themselves, and divide risks and profits and losses. It is extremely rare, consequently, that any one dealer or any one person buys at the lowest price of a season and sells at the highest price, making the utmost possible gain; or that any one buys at the highest and sells at the lowest price, incurring the maximum loss. Every dealer has losses as well as gains. On the whole, if he is shrewd and experienced, he gains more than he loses. He may lose money in one season but he will make money in another, and in the long run he will earn something in the nature of a professional income.

§ 3. When commodities are produced on a large scale for distant markets and for scattered purchasers, and middlemen become necessary links in the division of labor, it is inevitable that the middlemen should arrange to be near each other for the convenient disposal of their business. A street corner may serve as a meeting place. Traders in one commodity will settle near each other in a given street; in every great city there are dry goods streets, hardware streets, boot and shoe and leather streets, and so on. When in a populous and thriving country commodities are produced in large quantities and are necessarily dealt in by many persons, an exchange is set up—a room or building where the traders meet at fixed hours. Rules are agreed on, governing and interpreting their transactions in such detail that enormous sales are effected by a nod of the head and are recorded on scraps of paper with a few figures and initials. The actual dealings on exchanges are often done by brokers only, who are middlemen for the middlemen. They act simply as agents, earn their living by a

commission (usually at a rate extraordinarily small) on sales and purchases, and buy or sell for any one who chooses to transact business thru them.

The smooth dispatch of business on exchanges is further assisted by the "standardizing" of the articles dealt in, that is, by grading and classifying them according to quality. This process puts an end to all disputes regarding the quality of the things contracted for. Thus grain is examined as it reaches the Chicago market by publicly appointed inspectors and is graded as being No. 1, No. 2, No. 3. Thereafter, when a purchaser has his wheat delivered to him, neither he nor his vendor need inquire further whether it is of the stipulated quality. Delivery of elevator receipts, certifying the grade, satisfies all contracts. Any article that is homogeneous in quality or is easily classified into distinct grades can thus be dealt in with the minimum of friction. Grain is the typical commodity of this sort. Cotton is similar to it, thru its evenness of quality. Wool, which varies remarkably, is much less susceptible of rapid speculative purchase and sale. Attempts have been made to standardize iron, and in England a system of semi-official grading exists under which large transactions in it are carried on, but in the United States and on the Continent this mode of dealing in iron has never come into considerable use.

§ 4. Against the advantages which professional speculative dealings bring are to be set serious evils. These evils are made possible and are enhanced by the very facilities which enable speculation to work out its good effects.

When once a commodity has been standardized, a new possibility opens: anybody and everybody can deal in it. Ordinarily he who buys an article must know something about it. He must be able to judge whether what is offered to him is good or bad in quality, worth more or less. But on an exchange where commodities are officially graded no such questions arise. Only price, present and future, need be considered. Any one can buy if he thinks the present price low or sell if he thinks it high. Such buying and selling are done on an enormous scale by large numbers of persons who do not possess or wish to possess the articles they buy or sell and whose only concern is to make a profit by taking advantage of

fluctuations in prices. They virtually bet on the future price of the commodities and gamble about it as men gamble on cards or on horse races. True, in form their dealings are like any others on the exchange. The brokers receive orders to buy and sell, and by the rules of the exchanges are held responsible for delivery of the goods at the stipulated time. They hold their customers in turn to this same responsibility. But, tho thus in form like any other dealings on the better-known exchanges—the cotton and grain exchanges, for example—the immense majority of the transactions have in view no *bona fide* business. The machinery which has been devised for the easy and rapid transaction of business is utilized for large scale gambling.

The large part which "the public" plays in speculative transactions in commodities makes necessary a qualification, certainly a more guarded statement, of what has been said about the stabilizing effect of speculation in goods. The thick-and-thin defenders of business as it is often say or imply that the enormous mass of transactions on the exchanges, by whomsoever conducted, all works for the best; there are no evils so far as the community at large is concerned, while there is the gain from the lessening of price fluctuations. But when account is taken of the whole, it becomes doubtful how far the supposed economic gain in the way of price stability is in fact secured. The sellers and buyers include not only the cool-headed and well-informed dealers who are qualified to gauge the market influences but a multitude of ignorant and excitable persons, alternately reckless and timorous. On *a priori* grounds one would suspect that the doings of such a miscellaneous crowd would be as likely to increase the amplitude and rapidity of price fluctuations as to decrease them. And, while specific evidence on the question is hard to get, such as there is leads at the least to a non-committal view.

When all is said, it becomes clear that much more time and effort are given to these doings than can be reckoned advantageously to the community. Very much must be judged to be unproductive—mere waste of effort. Of course merchants, dealers, brokers are useful. The fact that some of their work, perhaps much of it, is not carried on as well as it might be does not show

that it is unproductive. In some discussion of this and kindred matters it is often alleged that there is "waste" because things are not so well done as they might be, which is quite different from saying that they are wasteful in the sense that no good at all comes from their being done. If the only persons who engaged in (say) the grain trade were such as systematically and continuously gave time and effort to it, their number would adjust itself automatically to the work required—in the same way as the number of carpenters or physicians adjusts itself to actual needs. Given the ways and traditions established at the time, the labor of all would be useful. But where there is "illegitimate" speculation on a great scale the labor involved seems to go beyond the range of what can be called productive. And not only the labor of the speculators but that of their agents is unproductive; the whole chain of operations adds nothing to the output of society. In no country is there so much of this parasitic activity as in the United States, for here all the conditions favorable to it are found—a highly developed division of labor, markets and exchanges on a great scale, and a population both venturesome and prosperous. "Business" to many an American means simply speculative gambling.

The "outside" speculators or the "public," like all amateur gamblers, are losers as a class and most of them are in the long run losers individually. The experienced professional dealers know better than they the probable course of prices, sell to them and buy from them to advantage, and on the whole make money from them. Occasionally a shrewd or lucky person makes a hit and carries off a large share of plunder from a successful operation on the exchange. This then acts on the imagination of others like a great prize won in a lottery. The chances that the speculative public will lose are almost as great as the chances that the purchasers of lottery tickets as a whole will lose: they amount virtually to a certainty.

Unmistakable as are the evils of gambling speculation, it is exceedingly difficult to check them by legislation, still more to put an end to them. The common law already makes void transactions which are sales in form merely and which contemplate a settlement only of the difference between present and future price. But

on the exchanges all transactions purport to be for the actual delivery of the commodities, and in strict legal effect are so. A remedial measure often suggested and sometimes tried is to prohibit buying and selling for future delivery, since it is in connection with such contracts that the gambling operations most often take place. But this would put an end also to the benefits which the community gets from contracts for futures, and it is a question whether the loss would not outweigh the gain. The common opinion of American and English economists is against the prohibition of future contracts. Yet the evils of speculative gambling are so great that something may be risked for the purpose of lessening them. Lotteries and avowed gambling houses have been prohibited, and all that can be done to stamp out other forms of gambling is welcome. No doubt the most effective remedy would be a better moral standard for all industry and an aroused public opinion against all kinds of gambling. But the worship of wealth, and the well-nigh universal desire to make money on easy terms even tho at the expense of others, together with the close association of this sort of speculation with business dealing accepted as legitimate, render it difficult to bring public opinion to bear.

§ 5. Stock exchange securities are ideally homogeneous and standardized. One share of a given corporation's stock is precisely as good as any other share. If it is easy for any one to buy grain or cotton even tho he has never looked at the articles, it is still easier for any one to buy stocks and bonds even tho he knows nothing about the corporation that issues them. At the same time, fluctuations in the prices of securities are large and frequent. Opinion regarding their probable course depends (or seems to depend) quite as much on general judgment and general prospects as on expert information. Hence rampant speculation by outsiders as well as by professionals. Here, as in the case of commodity speculation, the "public" loses in the immense majority of transactions. The professional speculators and dealers get the advantage of the miscellaneous public, both because they are better informed regarding the real prospects of the enterprises whose securities are dealt in and because they are (by a process of quasi-natural selection) persons shrewd in judging human nature and quick to take advantage

of the irresolute. Yet notwithstanding the constant losses, an unfailing stream of persons take fliers on the stock exchanges. There are probably few Americans of the well-to-do classes who have not at one time or another tried their hands at a stock speculation, and there are a great many who habitually gamble in stocks. The immense majority of these dealings are concentrated at the New York Stock Exchange, which is at once the greatest institution in the world for facilitating investment and the greatest of gambling hells.

The evil from the situation does not arise only or chiefly from the losses of the unsuccessful speculators. What these lose others gain, and usually there is not much to choose between winners and losers. But there is a net economic loss from the waste of much brains and energy on unproductive doings. The waste is more than that of the labor given directly, by brokers and their understrappers and by the speculators themselves. It is increased by the demoralization of many men in the community who take no great direct share in speculation. Like all gambling it saps character and distracts from the sober continuous work on which the common welfare rests.

There are economic losses in another direction, more serious and reaching farther. Widespread speculation, especially in stocks, tends to increase the intensity of industrial fluctuations and crises. Commodity speculation has similar effects but in a different way and in a different direction, and not so serious. As regards either class of transaction, it would be going altogether too far to say that we find here the root cause of fluctuations and crises. The causes are many and go deeper; speculation is a symptom and result rather than the main underlying thing. But an inflaming and intensifying factor it is and in so far must be reckoned among the causes. Gambling with dice or cards or on horse races may do no great harm outside the circle of the gamblers themselves. The same is probably true of gambling in cotton or grain. But widespread and furious speculation in securities links itself closely with the other causes of cycles in industry, with misdirected investment and fraudulent promotion, with booms and depressions, and so with those aspects of the modern industrial system which are most perplexing and also most ominous.

It must be frankly confessed that no really promising remedies have been suggested. Some excrescences have been aimed at in proposals for reform in New York—proposals which looked to improvement and no doubt did bring some improvement thru the revision and enforcement of the rules made by the exchanges for themselves. Rigging of the market, "wash sales," manipulation of prices with intent to deceive, are to be thus prevented. But even if all such tricks were cut out, the main cause of evil would remain. What brings the ill economic consequences is not so much these disreputable operations as the unbridled gambling in stocks, on a huge scale and in accordance with the rules of the game, by the well-informed and the half-informed and the entirely ignorant. In Germany a remedy tried in the early years of the twentieth century was the requirement of publicity in stock dealings thru enrollment of names and transactions on a register open to general inspection. It was expected that men would refrain from stock gambling, as they will from many doings of doubtful aspect, if they must be in the open. Such a requirement would be met in the United States by the objection that it intrudes on the sacrosanct secrecy of business—an objection commonly brought against public supervision of every sort and in itself of little weight. Much more serious is the failure of the device; in Germany the regulation had little effect, stock speculation remaining much the same in character and volume. A somewhat different line of repressive action was followed in the United States Act of 1934 for the regulation of stock exchanges. Here one of the main things directly aimed at—and rightly—was speculation by "the public" and especially by persons of small or modest means. The way taken for checking it was that brokers should be compelled to have larger margins from their customers. What success can be secured from this and other provisions of law is impossible of prediction; stock speculation may still remain much the same in character and volume. The problem is but one of many that we face in the present economic and social system, and indeed but one in the special group of problems which arise from large-scale production and corporate organization. It is among the things to be put in the balance when we try to weigh the advantages and disadvan-

tages of private property and free exchange. In some part the evils can be ended or greatly lessened. In some part they are likely to persist as long as private property persists, to be acquiesced in only if they are to be reckoned among the inevitable offsets to the individualist system.

CHAPTER 12

VALUE UNDER CONSTANT COST AND FREE COMPETITION

§ 1. The simplest case first assumed: a supply absolutely flexible, free competition, constant cost. Value then determined by cost.—§ 2. Illustration by diagram.—§ 3. The proposition points to a tendency or approximation only; to what happens in a "static," not in a "dynamic," state.—§ 4. Some explanations and qualifications. Flexibility in supply never perfect, often much impeded. Changes in demand from fashion. Good will. A small surplus above cost price may mean large profits.—§ 5. The extent and the importance of the divergence from competitive conditions.

§ 1. In the preceding chapter the problem was the adjustment of value when goods flow to market from an existing stock. Hence the supposition that the available supply was fixed: fixed not indeed for the day or the week, nor rigidly over any length of time, but fixed on the whole for the season or the period of production. But in most cases there is a considerable and sometimes rapid flexibility in supply. The amount produced and then put on the market changes. In what way do the variations in supply take place, and in what way do they affect the value of commodities?

We may begin by taking the simplest case and, for the purpose of bringing into sharp relief a principle, make again an extreme supposition. In the preceding discussion of demand and supply and of market value an absolutely fixed supply was assumed at the outset. Let now the other extreme be assumed, a supply absolutely flexible. Suppose a commodity produced, under the simplest conditions, by a large number of persons. Suppose that all these persons are competing with each other, that any one of them can easily engage in producing the commodity and as easily withdraw from producing it. Suppose all to be carrying on operations under the same conditions, no one of them producing more cheaply than another. Such a commodity would be brought to market under conditions of constant cost, and would be sold at a price conforming to that cost.[1] At any moment its value would

[1] By "constant cost" is meant not only that cost is uniform, but that it remains the same whether the total output be larger or smaller.

indeed be determined directly by its quantity—that is, by marginal vendibility as analyzed in the last three chapters. But if its value so determined were greater than its cost, more persons would be led to engage in its production, supply would increase and value would fall. If its value at any time were less than its cost, some persons would withdraw from its production, supply would decrease and value would rise. The greater the ease of entering on the industry and of withdrawing from it, the more rapid and certain would be the adjustment of supply to that amount which would just sell at cost price. If perfect flexibility in supply be assumed, the adjustment of value to cost would be perfect and the article would always sell for just what it cost to produce it.

Before proceeding further a word of explanation, and in some ways of warning, is needed as to the sense in which cost of production is here spoken of. The term is used in very nearly the ordinary commercial sense; it refers to the outlays which an employing capitalist must make in order to get a commodity to market. Chief among those is the outlay for the wages. Charges for material are another item. These charges, it is true, commonly imply that another capitalist has previously paid laborers to make the materials, which then have been sold to the particular employer in question; hence the latter may be said to have indirectly hired these other laborers also. Not only the wages paid to workmen, directly or indirectly, must be included but a reasonable remuneration for the employer's own time and trouble. This remuneration, like that of the workmen employed, is to be reckoned according to current market standards—what a workman or an employer of this kind would ordinarily receive for his labor. Again, interest on the capital used is to be included, reckoned also according to the current market rate. If the employer borrows the capital, he must pay the current rate of interest on it. If he owns his capital, he considers that he could get a return on it at that rate by lending it out to some one else, and he regards interest on his own capital precisely as he regards remuneration for his own labor—something for which a return at the usual rate is to be expected. It will be noticed that rent paid for land is not included in this enumeration, altho a business man would include it in his

reckoning of cost. The reasons for this omission will be explained when the subject of rent comes up for consideration.

These various outlays, or equivalents of outlay, are sometimes spoken of as "expenses of production." That term is sometimes used by way of distinction from "cost of production," emphasis being thus laid on the fact that the employing capitalist is concerned solely with what he pays for labor, for materials, for the use of free or fixed capital. When on the other hand the term "cost of production" is used with the design of pointing to a distinction from "expenses of production," reference is to what is also called by contrast "real" cost: to the labor of the hired workman, and not to his wages; to the trouble, anxiety and work of superintendence of the employer, not to his profits or ordinary gains; to the previous saving by which the capital has been accumulated, not to the interest on that capital. As will be seen at a later stage some of the most important and difficult problems of economics connect themselves with the distinction between cost of production in the sense of labor and sacrifice, and expenses of production in the sense of outlays.[1] For the present we need not do more than note the distinction in order to make clear in what sense we are speaking of cost. We mean by it not cost in the sense of labor and sacrifice but expenses of production—outlays of a capitalist. If we should think of a workman, or set of workmen, producing independently and without being hired by employers, we should reckon their "cost of production" for the purposes of the present discussion not in terms of hours or days of work (i.e. sacrifice) but in terms of the wages they would ordinarily get for their work.

§ 2. The mode in which value would be adjusted under the conditions of constant cost and absolutely flexible supply is indicated on Figure 5. The cost of the commodity is indicated by SO, the distance from the horizontal axis OX to the line SS'. Whatever the amount of the commodity produced, that cost remains the same for each unit brought to market; whether the quantity be OA, OB, OC, the cost per unit is the same. Hence SS', indicating the conditions of supply, runs parallel to OX. Let the line DD' indicate the conditions of demand, as in pre-

[1] See Chapter 53, Vol. II.

vious diagrams. It descends as quantity becomes greater, price falling with the increase in supply and the consequent lessening of vendibility. The supply of the commodity would then settle at the amount *OB* or *SB'*. The demand and supply lines would intersect at the point *B'*; there would be equilibrium at the quantity *OB* and the price *BB'* ($= SO$). If the supply should diminish to *OA*, the price might rise temporarily to *AA'*, *A'* being

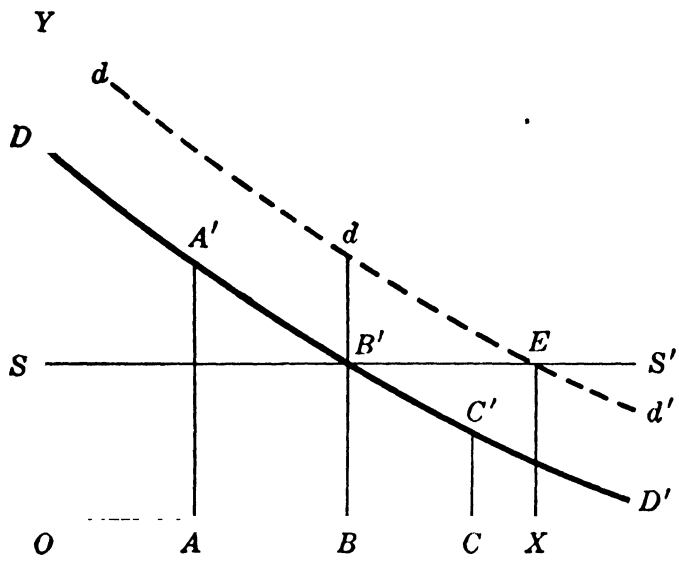

FIG. 5

the point at which the supply *OA* intersects the demand line. The marginal vendibility of the diminished supply would be raised to *AA'*; the smaller supply (*OA*) would sell at a higher price. But that higher price would lead under the conditions of constant cost to a prompt increase in supply. Producers would be getting more than sufficed to induce them to bring the commodity to market. They would compete with each other, increase supply, and so bring down price. If the supply should be increased not only to *B* but to *C*, the total being then *OC*, they would overreach themselves. For the amount *OC* the price would be *CC'*, the point of intersection with the demand line being then *C'*. This sum (*CC'*) is less than cost; some producers would promptly withdraw; supply would again diminish. For the quantity *OB* the price is just sufficient to make production worth while to all, and at that amount the supply would settle.

If now for any reason demand should increase, quantity would so increase as still to leave price at the same point. Suppose a change in fashion or other cause leading to an increased demand. This is represented by a shifting of the demand curve to the right. It is now *dd'*, whereas before it was *DD'*; at each several price more of the commodity is demanded than was demanded before at that price, and the marginal vendibility of any given supply is greater than it was before. With the supply *OB*, the price under these new conditions of demand would be not *BB'* but *BB'd*—higher than cost. Supply would again increase, until the total supply was *OX*. Then the demand line would be intersected at the point *E* and price would be $XE = BB'$. A new equilibrium would be established, not with a change in price but with a change in quantity supplied.

Under the conditions of constant cost and free competition, demand or marginal vendibility determine not price but quantity supplied. The proximate condition determining value is indeed always marginal vendibility, and where supply is fixed price is settled once for all by marginal vendibility. But where cost is constant and supply is completely flexible price cannot depart far from the level fixed by cost. The supply on the market will be such as can be disposed of at the cost price.

§ 3. The assumptions made at the beginning of this chapter—constant cost, flexible supply, free competition—are never, in a literal sense, in conformity with the facts of industry. There never is a case when these conditions are exactly fulfilled. None the less there is a wide range of industry in which an approximation toward their fulfillment is found and in which the principle of value under constant cost explains the broad facts.

Cost is never exactly equal for all producers. In the previous exposition such equality has been assumed. The fact that a general change in cost may take place, affecting all producers, does not necessarily alter the application of the reasoning. An invention or improvement may lower cost for all; the horizontal supply line on the diagram may be lowered; but the result is merely adjustment to a new level, not the introduction of a new set of conditions. If, however, the lowering of cost takes place not at the

same time for all the producers, nor in equal degrees, we have a new principle and a different case—production at varying cost. This is what in fact happens when inventions bring about a reduction in cost. The change takes place by successive steps. The more shrewd and enterprising of the competitors introduce the improvements first; others follow suit; gradually all adopt it. And by the time all have adopted one improvement another may be introduced, and the same steps are again gone thru. If there be a succession of changes—and such are likely in the highly progressive modern industries—equality of cost never exists. There are always some producers who are turning out their goods at lower cost than others.

None the less, over a great part of the industrial field there is a tendency to equality of cost. The differences between producers in cost are not permanent; the process is merely one of gradual and irregular adjustment to a new level.

Some writers have stated the difference between actual conditions and long-run tendencies by distinguishing between a static and a dynamic state. In the static state competition has worked out its full result, and unless there are permanent causes of variation commodities of the class here considered are produced at a uniform cost and always sold at a price corresponding precisely to that cost. In the dynamic state there is flux and change, variation in cost, oscillation of price. Yet the dynamic state tends to subside into the static. Its continuance depends on the incessant reappearance of disturbing forces.

The real problem is thus not whether price is in strict conformity to a cost of production uniform for all competitors but whether there is rough approximation to this situation and a tendency toward its full attainment in a static state. And such a tendency, to repeat, exists over a large part, perhaps the larger part, of the field of industry. A comparison has often been made to the tendency of the ocean to keep its level. Tides, currents, storms, cause disturbances, and it is never true in a literal sense that the level is maintained; none the less there is a normal level, and the actual height of the water tends to conform to it. Or a comparison might be made to the tendency of the air to maintain a certain

pressure. This pressure (measured by the barometer) is normally 29.9 inches at sea level. In fact it may be more or less, and rarely does the barometer stand precisely at the normal figure. None the less it oscillates about that figure, and tends to return to it. At any given height above sea level there will again be oscillations, with a different range and with a tendency to return to the new figure normal for that height.

§ 4. By way of illustration and explanation some of the disturbing causes may be briefly considered.

Most universal, perhaps, is lack of flexibility in supply. There never is complete ease of variation, such as to bring about the unfailing accommodation of supply to the precise quantity which will sell at the cost price. Even under the simplest conditions of handicraft production there is no such flexibility. There is less as plant and machinery become more important and every considerable change in output involves time and expense. Tho there is some flexibility in the output from an existing plant, it does not go far; any considerable increase in supply involves the making of new plant, and any considerable decrease involves the abandonment of some of the old. Changes of this sort, involving a readjustment of the preliminary investment, not only take place slowly but are much affected by vague general sentiment. Business men, not much less than others, go with the crowd. When the belief gets abroad that such and such an industry is "a good thing" they flock into it with no very careful calculation. On the other hand, when affairs go ill it is with reluctance that existing plants shut down. When the signs of increasing demand show themselves new plants are at first constructed slowly and hesitatingly; then, at the later stages of a sustained increase, with uncalculating excess. Here is one salient cause of the oscillations of modern industry, often affecting many trades at once and going far to explain industrial crises.

The prices of things subject to rapid changes in demand are especially fluctuating, even tho they be produced under conditions approximating those of constant cost. Almost all textile goods that are used for outer garments are affected by the caprices of fashion. For textiles worn by women the changes in demand are ex-

traordinary. The fabric which for the moment has come into fashion cannot be turned out as fast as the women want it, while that which was in fashion but a year ago can hardly be sold at any price. To such sharp changes in demand supply cannot be easily accommodated and the conformity of price to cost works itself out only as a rough sort of average.

§ 5. A couple of generations ago it was supposed that the prevailing case in the phenomena of value was the simple one of constant cost. As the twentieth century went on it became clear that monopoly and combination, once thought to be on the way to extinction, were supplanting free competition over wider and wider parts of the field. How completely it has been supplanted is not easy to say. An extreme view is that monopoly dominates the field once for all. This is an exaggeration made plausible by the conspicuousness of monopolistic conditions—an approach to monopoly even tho not full attainment—in some of the huge familiar instances like steel and oil. The exaggeration has been supported too by the unceasing endeavor to establish combinations, more or less monopolistic, in many other large-scale industries. It ignores the persistence of effective competition in a great mass of unspectacular industries, and ignores also the failures of the endeavors to secure effective combination in most of the industries where it has been attempted, failures as repeated as are the endeavors. While it is impossible to say how long these mixed conditions will persist, competition still prevails over a very large range, especially in the industries which directly supply the goods commonly bought by consumers. Such are not only the wholesale and retail trading concerns but the factories which are engaged in those stages of production which are nearest to the consumer—the flour mills, the manufacture of clothing, of boots and shoes, of household furniture and household appliances. True, in these also we see the tendency to large-scale production, and cannot but speculate whether it will proceed apace to quasi-monopolistic condition or finally to full-fledged monopoly.

The questions suggested by this possibility, already touched on in the preceding Book, will recur again and again; they bear not only on the organization of production and the theory of value but

on the functions of government, on socialism and private property, and so on the inmost meaning of man's welfare and happiness. As regards the present subject—the theory of value—it is enough to say that there is a wide range within which value is dominated, subject to all the qualifications just set forth and others to come, by competition under constant cost. True, the field of effective competition is probably becoming less; tho the belief in its shrinkage may be the result merely of our inadequate knowledge of the past. Research into the economic history of the earlier centuries of the modern period—before the nineteenth century—indicate that there was more of monopoly or semi-monopoly in those days than had been supposed. The larger the field in which competition remains free and active the more likely is it that the system of private property will endure, while its progressive diminution would tend to gradually undermine that system.

CHAPTER 13

VALUE AND VARYING COSTS. DIMINISHING RETURNS

§ 1. The equilibrium of value where marginal vendibility and marginal cost balance. The simile of the scissors.—§ 2. Permanent variations in cost affect long-run value differently from temporary variations.—§ 3. Diminishing returns.—§ 4. Permanent variations, or diminishing returns, appear most in the extractive industries.

§ 1. LET us suppose now that the several producers who compete with each other in putting a given article on the market have not the same facilities, that for some of them the costs of production are greater than for others. We need not concern our-

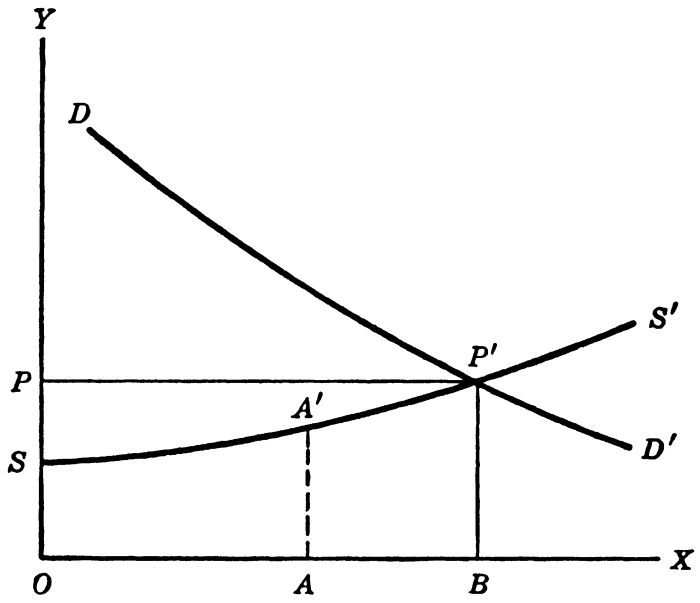

FIG. 6

selves for the present with the question *why* there are such differences. Let them be assumed to exist.

The situation is illustrated by the figure. The conditions of demand are again indicated by the descending line *DD'*. The

conditions of supply are indicated by the rising line SS'. The varying distance from the horizontal axis OX to the line SS' measures the varying cost of different installments of the supply. Some producers—those most favorably equipped—can put the commodity on the market at the comparatively low cost OS. Perhaps a moderate quantity can be so produced at constant cost. If the conditions of demand were such that only this moderate quantity were wanted at the constant cost price—if the demand curve were to intersect the supply curve not far from S—the normal price would be OS. So far the case would be identical with that studied in the preceding chapter. But the conditions of demand indicated by the line DD' are such that a much greater quantity is wanted at the price OS than can be furnished at that price. The supply put on the market increases, but as it increases additional installments can no longer be produced at the cost OS. With the quantity OA, for example, the cost of the last installment reaches AA'. As more is produced cost still increases, indicated by the continuing ascent of the supply curve from A' to P'. At P' finally the demand curve is met. At the price BP' $(= OP)$ the quantity OB can be disposed of. Equilibrium is established; the quantity demanded equals the quantity supplied; and price settles at the amount BP'.

The whole supply will be sold at the price OP $(= BP')$; and the selling value of the whole, i.e. the quantity multiplied by the price, will be indicated by the rectangle $OPP'B$. It is true that the more fortunate producers could sell the commodity to advantage at a less price. At the price OS or AA' they would still find it worth while to bring it to market. But the total quantity which will meet the demand at an equilibrium price cannot be supplied unless producers less fortunate contribute their quota. These will not do so unless they get their higher cost price BP'. At that price the whole supply will be disposed of. The more favorably situated producers will get the price necessary to induce their rivals, who have poorer facilities, to contribute to the supply.

We may speak of the producers at B, whose cost of production is BP', as the marginal producers. Their cost price is also the measure of the marginal vendibility of the commodity. Mar-

ginal cost and marginal vendibility thus coincide, and when they coincide there is equilibrium. If the quantity supplied should increase beyond *B*, in the direction of *X*, marginal vendibility would be less and marginal cost would be greater. Supply could not long be maintained beyond the point *B*, for producers would then be receiving less than cost. So long as the conditions of demand and supply remained as indicated by the lines *DD'* or *SS'*, price would settle at the amount *BP'*.

The relation of demand and supply to value is somewhat different here from what it was in the cases discussed in the preceding chapters. Where the supply of a commodity is fixed (the case which underlies the reasoning of Chapter 10), the value of a commodity is settled by the conditions of demand, that is, by the marginal vendibility of that supply. Where, on the other hand, the cost of a freely produced commodity is constant (the case discussed in Chapter 12), the value of the commodity is settled in the long run by the conditions of supply, by cost; demand in this case determines only the quantity which shall be put on the market. But in the case now under consideration the conditions of demand and of supply both have a permanent influence in settling price. As the quantity shifts, not only does marginal vendibility vary but marginal cost also. A decline of demand would not only lessen the quantity put on the market but would also lessen marginal cost. Conversely an increase of demand would not only cause more to be put on the market but would also raise normal price, since the additional quantity would be produced at greater cost. Hence demand and supply—marginal vendibility and cost—mutually determine normal price.

Professor Marshall has ingeniously compared the influence of demand and supply to the working of a pair of scissors. If one blade of a pair of scissors is held still, and the other moves, we may say that the second does the cutting. Yet it could not cut unless the other blade were there. So when supply is fixed, we may say that demand settles value, yet it does so only because supply is there and does not move. When cost is constant, we may say that cost settles value. Yet it does so only because there is a demand for the commodity, and because supply adjusts

itself to the amount which will be demanded at the cost price. If cost is variable in the manner discussed in the present chapter, both supply and demand—both cost and vendibility—exercise a mutual influence on normal price. Both blades of the scissors are in motion. All the various manifestations of value can be analyzed as interactions of supply and demand. Neither can be said to settle value independently of the other

§ 2. The differences in advantage between producers may arise from permanent or from temporary causes. According as the differences are temporary or permanent they are of very different significance for the theory of value and for the welfare of society.

Differences of a temporary sort are the most common. They are so common that they may be said in one sense to be universal. As indicated in the last chapter, it probably never happens in communities familiar to us that all those engaged in a given industry are carrying on their operations in the same way. Some having better plant, better organization, better location than others can bring their products to market at less expense and, selling at the same price, can reap larger gains.

But these differences, if their causes are not permanent, tend constantly to disappear. If one man has better plant or machinery than another, and if there be no permanent reason why the second should not also set up the better outfit, he is likely sooner or later to do so. If he does not he is likely to be driven out of the market. Others who adopt the more effective method of production will increase the quantity they put on the market, and will be able to undersell him without foregoing a profit. Where the methods of cheapened production are open to all they will sooner or later be adopted by all.

We say sooner or later, for the process takes time, especially when changes in the arts are rapid. The civilized world has been for generations in a dynamic state. Causes of differences are constantly appearing, disappearing and reappearing. At any given time the usual conditions are those not of uniform cost but of varying cost.

Under these conditions value cannot be said to be determined by marginal cost of production. Proximately value is always de-

termined by the marginal vendibility of the supply. Given the total supply that comes on the market—whether offered in large sudden doses or by gradual increments—and the price will be such that the whole is sold. For the marginal producer this price may or may not be equal at any given time to his cost. With the oscillations of demand and the various other causes of non-adjustment to normal conditions which were considered in the preceding chapter the price for the time being may be such as to make the marginal producer prosperous or such as to make him a bankrupt. If he happens to be prosperous his more enterprising and successful rivals, the infra-marginal producers, are even more so and are tempted to extend their operations. If he is on the way to bankruptcy they may yet be able to hold their own. In time he disappears and his better-equipped or better-situated rivals supplant him. In time, then, it is cost of production at their hands which acts on supply and thus acts on price. In other words, disregarding temporary and seasonal fluctuations, the principle of constant cost regulates long-run value where there are non-permanent differences between rival producers. It is cost of production at the hands of the capable and well-equipped producers that settles the long-run price as distinguished from the temporary price.[1]

The outcome is different where permanent causes underlie the varying costs of the several producers. Then cost at the hands of the marginal producer does settle the long-run price. The point about which oscillations range and to which price tends to conform is cost for the least advantageous producer. Without him the total supply cannot be enlarged to the point at which there is an equilibrium of normal supply and demand. If indeed there were no limit to the amount which the more advantageous producers could bring to market—if this fortunate set could increase the output indefinitely at constant cost—the marginal producer would be driven out and the conditions would be those of constant cost. There being such a limit, he must be called on for the maintenance of supply and there must be in the long run a price which will make it worth his while to contribute. Value is then

[1] Compare what is said in Chapter 49, §§ 1, 2, Vol. II.

determined by cost to the marginal producer; but at what point in the varying scale of costs that producer will be depends on the conditions of demand.

§ 3. Instead of speaking of varying cost or increasing cost we may speak of diminishing returns.

Increasing cost and diminishing returns are opposite aspects of the same tendency. Looking again at the diagram we may see that the marginal producer at *B* has, for the same addition to the supply, greater expenses than the better-situated producers at *A* and *O*. As the quantity put on the market increases along the axis *OX*, cost for every fresh installment becomes greater. With every proportional increase in outlay there is a tendency to diminishing return. It matters not whether we say that the tendency to diminishing return is experienced by the infra-marginal producers themselves or by those producers whom we have regarded as marginal. Unless it were experienced by the former the particular situation considered in this chapter—that, namely, of variations in cost due to permanent causes—would not exist at all. If these better situated persons could enlarge their several contributions to the output indefinitely at the same cost, they would supply at that constant cost the whole quantity demanded. No interlopers could then get in. But each and all of them would find, as additions are made to the supply, that a stage arises where the addition cannot be made on the same easy terms as for the earlier installments. Diminishing returns are encountered. New producers then may come in, but these in turn encounter the same difficulty. Some increments of supply they may produce on comparatively easy terms, but if they try to add more the terms become less easy for them too. All along the line there is an increase of cost as more is produced, and price settles at the point of cost for the last increment.

Something more is to be said about the way in which the distinction between real cost and money cost bears on this question. In the preceding, the more advantageous sources of supply have been assumed to be owned and used by the same person; it is the owner of the land who uses it and disposes of the product. Suppose now he leases the land to a tenant. To the tenant the sum he

pays in the way of a rent is an expense like any other, and is reckoned by him among his money costs of production. But it is a money cost which arises in a way of its own. It does not serve to pay one who undergoes real cost; it is not like wages paid to a laborer; it arises because the owner of the land possesses a superior instrument. That instrument would yield *him* an unusual or extra return if he applied the labor himself. If then he lets the instrument to another, he will ask and will readily get income to the amount of this extra return. The tenant neither loses nor gains by the payment, which simply puts him in the same position as others who pay less rent (for good land, but not so good as the best) or no rent at all (for land just on the margin). Nor does the rent paid, large or small, cause the price of the product to be higher: that is determined, in the manner just described, by the cost of the marginal product—that which has necessitated the largest application of labor and capital per unit of output. To the individual producer the rent paid is a money cost like any other of his outlays but as regards the price of the product and as regards the welfare of the community, it is a money cost which has a different meaning and a different effect from those money costs which are incurred because there are real costs back of them.

This is the classic theory of rent, stated in its baldest terms. It needs both qualification and explanation; on the other hand it also has wider applications than was supposed when first worked out. More will be said, as we proceed, about its meaning for various parts of the field of economics. Properly qualified, explained, understood, it remains an abiding and essential tool of analysis.

§ 4. In what circumstances, and over how great a range of industries, do we find varying cost or diminishing returns? In general, differences in cost are permanent in the extractive industries—in agriculture, forestry, mining.

In agriculture, good land yields more to labor than land less good. The prairies of Illinois are more fertile than the stony fields of New England, the black earth of central Russia than the sandy soil of Brandenburg. All the climatic factors—such as sunshine, precipitation, the length of the seasons—have their influence, as well as the physical and chemical constitution of the soil.

Not only are there such differences but there is further an unmistakable tendency to diminishing returns on any one plot of land. The amount of produce which can be obtained from the best land is limited; and the amount which can be obtained from that land under the best conditions is limited. By applying more labor and capital, it is usually possible to add to the produce from a given piece of land, but it is not possible to get more produce in proportion to the addition of labor and capital. Hence there are permanent differences, not merely between different soils but between the successive applications of labor and capital on the same soil. Agricultural production, much as it has been modified during the last century, still presents typically the application of the principle of value which we are now considering.

In forests, likewise, there are differences of the same sorts. Some are better than others. Advantage in location and accessibility plays no less a part than advantage in the size and character of the timber, yet either kind of advantage counts. Mines present differences of an analogous kind; they are affected both by accessibility to the market and intrinsic productiveness. Both forests and mines have industrial peculiarities, especially in their development during very modern times; but as regards both, the general conditions of varying cost and diminishing returns hold good.

In manufacturing industries, which shape and transform the materials brought out by the extractive industries, the principle of diminishing returns is applicable in less degree. The differences in cost between competing producers are commonly of the transitional sort. But sometimes they have permanent causes. One manufacturer may have an unequaled site on a harbor front or more water power than others. If he owns a site commanding such power, he can produce with less labor and sacrifice (real cost) than those not so favored; if he has it by lease from the owner, his rent is one of his money costs, and is to be analyzed in the way indicated in the preceding paragraph. In the earlier days of the development of power and machinery, a first-rate water power was of great advantage. Later steam largely superseded water power, partly because of the great advances in the efficiency and economy of steam engines, partly because they could be set up at any de-

sired place and so permitted better access to markets or to materials. In recent years the generation and transmission of electric power has again made falling water important, and may prove the cause of enduring differences in the effectiveness of manufacturing establishments.

CHAPTER 14

VALUE AND INCREASING RETURNS

§ 1. External economies.—§ 2. Internal economies. Large-scale establishments. —§ 3. Effects of the two on industry at large, on the development of monopoly.—§ 4. What is meant by a "law" of increasing returns. How far there is a law as regards external economies, as regards internal economies. —§§ 5, 6. A theoretical treatment of increasing returns. Successive or historical supply curves. The interaction, over long periods, of increasing and diminishing returns.

§ 1. THE topics to be taken up in this chapter can best be discussed by introducing first the distinction which has been drawn by Marshall between external and internal economies.[1]

First, external economies. The term "external" implies that they arise outside of the establishment which applies them. Their occurrence commonly results from enlargement of an industry's total output. If we are to speak of a law of increasing returns from external economies, what it means is that there arises, not merely as a frequent occurrence but as a steady and predictable consequence, a gain in effectiveness as total output enlarges—larger product per unit of cost, lower price per unit of product. An example at once simple and typical is the diminution in cost of spindles, looms, and the like which takes place as these are made in separate establishments specializing in their manufacture. The more cotton mills there are, and the larger the number of these machines called for, the cheaper they can be made; and as the machinery becomes cheaper the expenses of the cotton manufacturer become less. Again, the construction of steel ships in the United States was long carried on for a much smaller total tonnage than in Great Britain. Consequently various accessory parts needed for ships—compasses, capstans, winches, donkey engines—were called for in much larger quantity in Great Britain, were systematically

[1] I am not sure that Marshall's definition of these terms is the same as underlies the discussions in these pages. Differences there are but they do not affect the main tenor of the chapter, and least of all imply any lack of appreciation for Marshall's path-breaking work.

and uniformly made on a large scale, and were cheaper for the shipbuilder.[1] Still another example is in the boot and shoe manufacture. When this is carried on with huge output, and especially when a number of establishments are in the same locality, subsidiary industries arise which supply cheaply the special tools, materials, and fittings—the shoestrings, eyes, metal fittings, lasts, boxes for packing, not to mention the machinery. The gain in external economies of this sort is one of the reasons for the concentration of an industry in a given place: of shoe manufacturing in Brockton and Lynn, Mass., of silk manufacturing in Paterson, N. J., of metal wares in Bridgeport, Conn. In every such place the factories, merely because of their number, command resources and economies which an isolated establishment finds it harder to secure.

All this is another way of stating the old doctrine of Adam Smith that the division of labor is "limited by the extent of the market." An enlargement of the market from the mere growth of population leads of itself to further elaboration of the division of labor.

A gain of this sort, less important yet considerable, comes from the presence of a large experienced labor force. In almost every single establishment the workmen are more or less shifting. The "labor turnover" is higher in industries exposed to seasonal fluctuations, as the boot and shoe manufacture is, or to irregularities of investment demand as in the case of establishments making machinery. It is lower where steady wants are supplied, as in the manufacture of soap, and where long-established businesses are conducted by firms of settled prestige. In many ways these shifts in the labor supply have ill consequences, yet seem to be an inevitable outcome not only of the variations in the demand for labor but of the psychology of factory labor. Certain it is that workmen come and go, and new men must be found to replace those who leave. Mechanics are more likely to be found in manufacturing centers, and especially in centers where there are industries

[1] These external economies would indeed have been at the disposal of the American shipbuilder if he could have brought the parts from Great Britain without restriction. But the United States imposed on them a heavy customs duty; hence, whether procured from abroad or at home, they were dearer.

similar in their general character, even tho not quite the same. In an isolated establishment the loss of a few skilled and trained workmen might cripple the whole. In an industry which has grown to considerable dimensions and which is concentrated in certain towns or districts, there is a general diffusion of skill in its various branches, the turnover of labor is at all times considerable, and gaps here and there can be filled with comparative ease. No doubt in such centers workmen are more likely to be organized in unions, and to press for higher wages, which from the employers' point of view may seem a disadvantage. But the fact that manufacturing towns grow shows that they offer net advantages.

§ 2. Internal economies are those which arise within the establishment itself. While they may not be completely independent of total output, they are not directly induced by the growth of the industry as a whole.

A characteristic example is "scientific management": the methodical study of the operations of an establishment with a view to learning exactly how best to use the machines and how best to handle the men. Another example is the moving platform introduced by Ford for assembling the several parts of the motor car. The idea was not a novel one, and obviously the device was not within the range of possibility for motor works until these had reached a large size. But once the adequate size was reached, the improvement came within the concern. Steam-engines and electric-power machinery for motor works, on the other hand, come from without, being designed and constructed by specialized plants for all sorts and kinds of industries; these are typical external economies. Not so the moving platform, which each motor concern puts up for itself.

The kind of internal economy that is most discussed by economists and is of greatest social significance is that of the size of the individual concern. The optimum size depends on the effectiveness of large-scale production as distinguished from that accruing from large output of the industry as a whole. Large-scale production depends primarily on the organization and management of the individual plant, on the division of labor within it and the marshaling of the successive steps. Given an industry with un-

changed total output, there may still occur an enlargement of the average size of the individual concerns, and so a less number of concerns. Such enlargements, if they do come, result from the schemes and experiments of the owners and managers, each trying to find the kind of outfit that turns out the product at the lowest cost per unit.

Connected with this is the matter of large-scale management—horizontal combination. This is no more than a particular phase of large-scale operation, conspicuous in the United States for half a century, common in Germany (tho rivaled in that country by the cartels, which do not usually contemplate united management), not unfamiliar elsewhere. It is not clear to what degree its growth has been actuated by the expectation of achieving internal economies, how far by that of securing a monopolistic or semi-monopolistic control of price, a topic on which something has been said already and more is still to be said.[1] It is enough at this point to note that horizontal combination has a place among possible internal economies, tho a dubious and uncertain one.

Vertical combination—the integration of industry—is to be classed with hardly any qualifications as an internal economy. It is a matter of the intercalation, extension, organization of a series of operations planned to fit into each other from bottom to top. As has already been said, this has been carried out in modern industry to an extent undreamed of before the nineteenth century. In many cases, moreover, it has resulted in making the individual concern enormous. The process in itself has rarely led to complete monopoly. Where this has occurred, the main cause has been control of a scarce raw material, as in the case of the aluminum industry. But integration on a large scale has entailed not infrequently the reduction of the number of individual concerns to the point where they are tempted to the other form, the trust or horizontal combination, in the expectation that still further economics will thereby be secured and one huge monopoly built up. The internal economies expected from horizontal combination have proved less easy of attainment than those from vertical.

[1] See Chapter 4 on Large Scale Production and Chapter 65, Vol. II, on Combinations.

The distinction between internal and external economies cannot be drawn sharply, nor is it necessary to draw it with sharpness. Improvements and inventions of every kind are affected by the stage of development at which the technical processes stand at the time, and in this sense all are partly external. All, too, prove on careful historical inquiry to have had a gradual growth and to have reached the stage of successful application thru trial in many scattered places. So it was with the moving platform. So it is with the practice of putting a large manufacturing establishment not in the heart of a city with high buildings but in the open country with low sheds spreading over large space. This practice has obviously been affected by the improvements in transportation and the consequent ease of moving men, materials, finished products. It may thus be said to come from without and to be in that sense external. Yet it cannot be standardized, nor adopted in any given case by merely turning to firms which specialize in construction of this kind. The case is different and the external influence is more direct with the highly intricate yet sharply standardized processes of the industries that make machines. The machine-tool industry, in which are made both small simple parts and large complex instruments for any and every sort of manufacturing plant; the great establishments which make textile machinery ready to put into place and to start operation at once; the similar establishments which make complete outfits for a shoe factory—these stand for something different. It needs but a moderate acquaintance with the characteristics of modern industry to make clear how far external economies of this kind have been carried, farthest of all in the United States.

§ 3. Consider some of the wider aspects of the two kinds of economies. Which of them has played the larger part in the enormous expansion of productive capacity during the nineteenth and twentieth centuries? It is not easy to say. The economies which are primarily internal are the more conspicuous. Often they are dramatic, such as Stephenson's discovery that the friction of the heavy locomotive on a flat rail was the key to the railway, or the Bessemer process for steel-making, or Edison's vacuum for the incandescent lamp. Less dramatic are the slowly accumulating ex-

ternal economies whose effects permeate every corner and cranny of a nation's productive equipment. All things considered, I should be disposed to say that by the close of the nineteenth century it was the external economies, first made possible and then promoted by the advances in transportation and communication, which have come to act most effectively toward enlarging productive capacity.

A different question, more easy to answer, is the effect of the two on the growth of monopoly and monopoly power. Clearly those internal economies operate toward monopoly which arise from the increase in the size of the individual plant. *If* each enlargement in the size of a concern means lower cost per unit of the product, then the larger concerns will be able to undersell the smaller and gradually to buy them out or push them out and in the end nothing will be left but the one huge concern. The form in which this sort of result would seem most likely to come is that of horizontal combination and large-scale management. *If* (emphasis again on the condition) this leads to additional economies without limit, the one great combine or trust will dominate. It may have a number of plants technologically identical but scattered over the land, but the fundamental industrial characteristic would be that of a lower range of costs secured within the total agglomeration. On this topic something has already been said and more will be said in later chapters.[1]

§ 4. We turn now to distinctions and discussions of a more theoretical kind.

"Increasing returns" may signify an observed fact; or it may signify a persisting tendency, a "law." As regards the theory of value, a "law" of increasing returns would mean that an enlargement of total output tends of itself to bring a lowering of costs per unit, and hence under competitive conditions a decline in price. To say there is such a law for an industry means that, given this one factor operating by itself, there is a predictable outcome. Changes whose occurrence and effect cannot be foreseen are beyond the law's range. Of course, these occur; such are the discovery of new and extraordinary mineral deposits, which has played so

[1] Especially Chapter 65, Vol. II, on Trusts and Combinations.

large a part in the development of modern industry. In a wider sense these changes are of course not haphazard; if the occurrence of the deposits seems to us a matter of chance, it is only because we cannot discern an order or sequence. For the present purpose, the question is whether we can discern a law of increasing returns in the sense just indicated: is larger output, operating by itself, followed by lower cost per unit? Suppose, for instance, that there is an increased demand for a given article and hence a larger output—nothing else; will the marginal cost fall? If this is to be expected from the mere fact of larger output, the industry conforms to the law of increasing returns.

As regards external economies, the case is comparatively simple. As output increases, economies of this kind are likely to be introduced. There is a law of increasing returns.

This generalization, to be sure, can be maintained only under the industrial conditions of the countries of western civilization, that is, with reference to those times and places with which economic science has been chiefly concerned. It could be applied but hesitatingly, if at all, to the economic life of the peoples of China and India. Even for the countries of western Europe its applicability is questionable until we reach the modern period. It has come to be of unmistakable significance with the enormous advances in transportation which began in the nineteenth century. The railway, the steamship, the motor; the post, the telegraph, the telephone, universal literacy—these have made possible a division of labor more far-reaching in effect than would otherwise have been possible, and so have led to external economies more and more pervasive. It is with reference to the total output in modern times that we can say with some confidence that its increase has brought a lessening of cost per unit in a great range of industries. So large has been the number of industries, so great their output, that the economic system as a whole has shown increasing returns.

As regards internal economies, the situation is not so simple. This is more especially the case with that internal economy which is of most significance for the social problems: the gain from large scale production. Can it be said that an increase of total output in

a given industry leads to lower costs, to increased returns, for the reason that total enlargement of itself tends to make the optimum size of the individual establishment greater?

No unqualified answer can be given. At first blush a negative seems to be called for. If the gains from enlarging a concern arise from greater effectiveness within, what does it matter whether the industry as a whole is expanding? Suppose an industry is declining and total output becoming smaller; will not changes in technology or management still make it possible to lessen cost and price by increasing the size of the individual firm? The number of firms will then be less, but each one will be more efficient. And conversely the mere fact of greater aggregate output would not seem of itself to stimulate technology and management in the direction of enlarging the concern. Suppose there is an increase of demand —simply this and nothing more. That means more output; then external economies, lower cost, lower price; then more firms in the industry—but not firms that are individually larger.

This sort of reasoning, however, cannot be applied offhand to a world that is moving and changing. It ignores the complexities and frictions of the modern world, the temper and the environment of the business men. A man who sees the way to a bigger and better scale of operations may be convinced that the change would work to advantage, yet when it involves a long period of struggle before the existing smaller concerns are crowded out he will hesitate. If, however, demand is increasing, if there is slack to be taken up, he will put in the larger outfit without hesitation. At such times it is in the air to go ahead with any promising scheme. Business men are not to be thought of as quasi-automatic agents, always cool and calculating, uninfluenced by the atmosphere prevailing in the world of affairs. The attainment all around of a new optimum thru enlargement of the size of the representative firm comes more quickly and surely in times of general buoyancy and expansion. And an element in the long-run situation has been for generations the steady advance of large-scale production as well as of total output—hesitation in times of depression, but buoyancy and a rapid pace in times of activity.

As a matter of logic, of "pure" theory, all this would seem to

bear merely on the *pace* at which the larger concern would displace the smaller, not on the eventual outcome. Even with a total output which is stationary the big concern, if its costs are the lower, will become the representative one in the end. It is hardly to be doubted that, given time enough, this would ensue. But time and temper count for much in the industrial development of modern times; and as regards internal economies as well as external, we have always to bear in mind that it is of modern times only that we can speak with any assurance. The prevailing industrial atmosphere, the business temper, a growth of population and a secular expansion of production—when such factors as these spur enterprise generation after generation, large-scale production is promoted not only more speedily but more effectively.

The same questionings arise about other ways in which returns have been increased. Are technological inventions to be regarded as internal or external economies? Sometimes they arise within a concern. Sometimes they seem to arise spontaneously from the speculations, dreamings, inborn bent of the isolated inventor. In either case they are not easy to classify under the rubrics here considered. But in the modern world they are sometimes deliberate, the outcome of research directed to a given problem in special laboratories. In this way they are often associated with the huge concern—a matter of internal economies. And yet again these may be in the main a mere following-up of clues already given, not anything radically new. All inventions have a long history, and perhaps all are largely induced—induced, that is, in the sense that devices suggest themselves to the inventive minds when demand increases and output enlarges, and in that sense they may be regarded as external.

Summing up, we may say that in modern times enlargement of the total output has been followed by a lessening of costs per unit of product in a wide range of industries. So wide has been that range as to make this outcome the typical one. As regards those economies which are most clearly of the external kind, being the predictable results of more intricate division of labor, they are not only typical but of wide range. Others not so easy of

analysis have also come more thick and fast as total output enlarges. In general, the march of modern industry has been toward increasing returns and thereby toward material progress all around.

§ 5. We turn now to some theoretical aspects of increasing returns which fit into the analysis made in the preceding chapters on constant costs and diminishing returns.

Suppose that there is an expansion of a given industry. Demand increases. The cause may be a change in tastes; people want the article more. Or it may be a mere increase of population; there are more people, and while each of them wants only as much as before, the total wanted at any given price is larger. In either of these cases the demand curve shifts to the right. In other cases the demand curve might not shift at all; provided merely that the demand is elastic, absorbing larger quantities at a lower price, increase of output will occur.

Suppose, as a first variant of the case, that the industry is one of constant, i.e. uniform costs. While uniformity of costs is maintained productivity all around is increased, say by some spontaneous ("autonomous") invention. A new process, equally at the command of all the producers, enables each one to bring the commodity to market at lower cost. Competition compels them all to sell at lower cost. The supply curve drops, but remains a horizontal line. More of the commodity is made and sold; and if the demand is highly elastic the total sales will increase greatly. It is not necessary that the demand curve should shift, we need suppose only that the demand is elastic.

Figure 7 illustrates this simple situation. $S'a$ is the supply curve at the start, Aa is the price. The improvement lowers the cost. The supply curve drops to $S''b$, and the price drops to Bb. And here another element enters. There being an increase of the total produced, there is occasion for the economies which are made possible by the enlargement of the output. External economies will certainly be promoted; internal economies very likely will also be. Suppose an alert community in which competition is free and in which the possibilities are utilized to the full. The supply curve falls still further; it goes down to S''', and for the purpose

of the present illustration may be supposed to remain there.[1]

The curve *abc* then may be called a supply curve, but it is one of a different kind from the more familiar one. It is a "successive" or "historical" curve. The supply curves (horizontal lines) *S'a*, *S''b*, *S'''c* may be called contemporaneous cost curves. They represent a series of costs incurred at one and the same time, and at any given time they are all the same: they are co-existing costs. The points *S'*, *S''*, *S'''* indicate costs (*Aa*, *Bb*, *Cc*) incurred at successive stages, the earlier ones being displaced by the later. When the second of these stages is reached, the cost *S''* is incurred by all, the cost *Aa* no longer exists. Similarly when the third stage is reached, the cost *S'''* is incurred by all, and *Bb* also has ceased to exist.

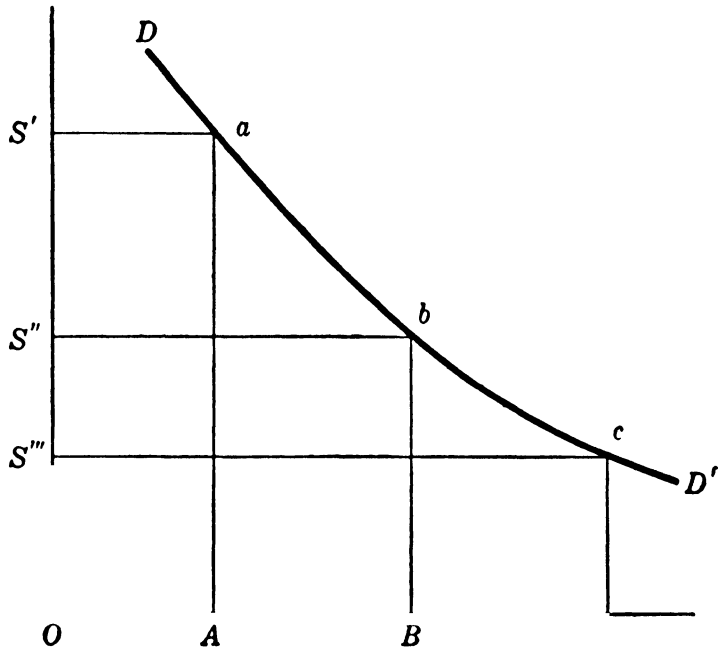

FIG. 7

It is evident that in this first simple variant of the case the historical or successive curve *abc* follows exactly the demand curve. It intersects the demand curve at different points one after another, each point of intersection indicating an equilibrium of supply and demand. So long as demand remains unchanged—so

[1] It will remain at this point if at the price *Cc* demand is satiated; that is, if any further decline in price would bring no increase in quantity demand. The demand curve would then fall perpendicularly from *c* to C.

long as the demand curve does not shift—the historical cost curve must always be identical with the demand curve, regardless of the character of the forces that bring about the decline in cost.

Suppose now that we have an increase of demand in the schedule sense; the demand curve shifts to the right. Then the historical cost curve is no longer coincident with the demand curve.

Figure 8 shows the new situation. The demand curve has shifted from $D'D'$ to $D''D''$, and then further to $D'''D'''$. At any

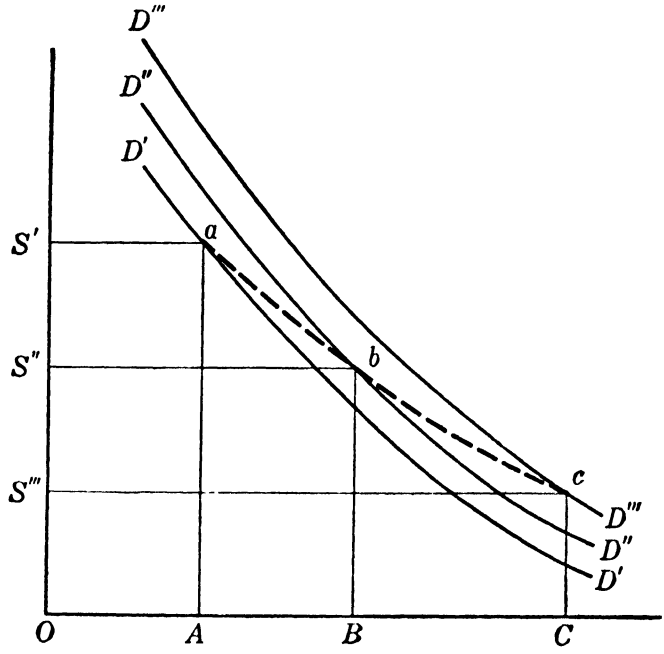

FIG. 8

given price more is demanded than before, and more is produced. The "law" of increasing returns being in operation, and the mere fact of greater output bringing a fall in cost, price falls. If the demand curve continues to shift to the right the output becomes still larger, the price falls still more. The process will stop (say) when the quantity OC has been reached and the price is Cc.

Connecting the three points of equilibrium abc, we have a historical (successive) supply curve which is quite distinct from any demand curve. It shows successive points of equilibrium, of which the earlier ones have disappeared as later ones are reached. Only the last one, at c, remains when the final and definitive equilibrium is reached.

§ 6. Suppose next we have the conditions not of uniform costs but of permanently varying costs, not constant returns but diminishing returns. At any one stage the supply curve has an upward inclination, indicating that some among the producers have low costs, others higher costs. A given outlay brings a larger return for some parts of the supply than for other parts.

Assume now an increase of demand, a shift of the demand curve to the right. Suppose the increase to be considerable (larger than that suggested visually on the chart); great enough to lead

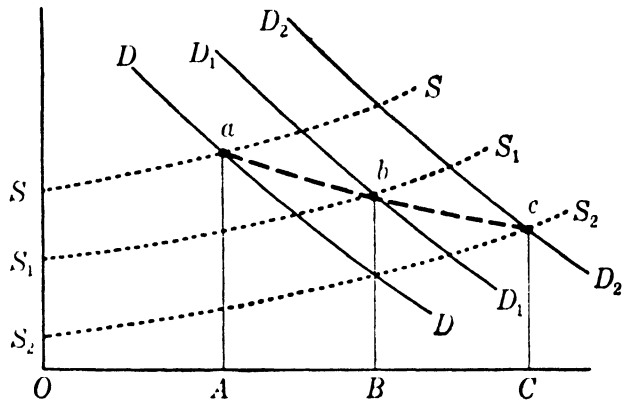

FIG. 9

to external economies, perhaps to internal economies as well. All producers will gain (not necessarily at once) by the new and more advantageous conditions. The supply curve, while lower than before, may or may not alter in shape; we may assume it to be in this respect unaltered. Three stages are represented on Figure 9. The demand curve, at first DD, shifts to the right, and is at D_1D_1; then shifts further to the right, and is at D_2D_2. The supply curve, at first SS, falls to S_1S_1 and then further to S_2S_2. At each of the successive stages there is a point of equilibrium where supply and demand are equal. Connect these three points *abc* and we have the historical cost curve.

The historical cost curve on this chart has a negative inclination, indicating that cost and price have steadily fallen thru the period. This is not necessarily the case. That curve may, like the contemporaneous curve of each of the three stages, have a positive inclination. The outcome in this regard depends on the extent

of the change in demand, the extent to which increasing output brings external and internal economies, and the character of the demand curve itself. In general, the greater the elasticity of demand the greater must be the drop in the cost curve in order to get a decreasing historical cost curve. The outcome depends too on the constancy of the conditions both of demand and supply; either curve may change its shape as well as its position. The particular conditions represented by the diagram have been selected not only because they are comparatively simple but because they bring out one great fact of fundamental importance in economic development during the nineteenth and twentieth centuries. This is the historical decline in cost and price of the commodities produced under conditions of costs that persistently vary—of diminishing returns.

"Diminishing returns" has an ominous ring; "increasing returns" has a cheerful one. The phrases seem to be antithetical, indicating situations that cannot exist together. But Figure 9 indicates that in the course of time there may be falling prices while supply curves are still positively inclined. It is for this reason that it often is better to speak of varying rather than of diminishing returns. Tho the conditions are for long periods, such that, costs and returns are not the same for all increments of supply, nevertheless costs may be falling per unit of output all along the line, and returns increasing. And this may be so even tho the causes that make for diminishing returns are permanently in operation. The *differences* between the several producers may persist while the cost level for each and every one is falling.

Reference has already been made to the fact that where evidence is available for making a cross-section of costs existing at a given time for any industry, we find without exception that costs are not uniform but varying. The phenomenon does not always have the same meaning. Often it arises from non-permanent causes, being the result of an uneven march of progress. In a period of general advance, some producers are always ahead while others are laggards. As time goes on, the laggards may drop out entirely or they may catch up. With the next advance the phenomenon then may repeat itself, but the high-cost concerns

at any one period disappear as such—either they adopt the better methods and lower their costs or they are forced out. Such is the typical course of events when there is progress in the manufacturing industries. It is otherwise in the extractive industries, and particularly in agriculture and forestry. Here, in any country long inhabited by the civilized peoples, the same land remains in use year after year. Parts of it are permanently better than other parts, and the better sites steadily produce more than the poorer. Improvements in methods of cultivation, in the maintenance of fertility, in agricultural machinery may be applied on all. But they do not wipe out the differences. The supply curve shows not only differences between the sources of supply but persisting differences.

Hence the general upward movement—increase of product with decline in price—has been greater and more unmistakable in the manufacturing industries than in the extractive. While it has appeared all around, and at times it has even seemed to be as marked in agriculture as elsewhere, on the whole man has become more adept at fashioning the stuff after it has been got from the earth than in extracting it. On this large subject something has been said already; more will be said in chapters to come.[1]

[1] See Vol. II, Chapter 46, Rent of Mines, etc.

CHAPTER 15

JOINT SUPPLY AND JOINT DEMAND

§ 1. Joint supply, or joint cost: effect of increase or decrease in demand. Influence of separable items of expense. "By-products." Complex case where both monopoly and joint cost exist. Influence of large plant.—§ 2. Joint demand. The constituent most limited in supply feels most the effect of an increase of demand. Labor in building trades as an illustration. Joint demand usually causes peculiarities less enduring than those arising from joint supply.

§ 1. Not infrequently commodities are produced at joint cost; the same operations which turn out one in the group turn out another also. Such are mutton and wool; beef, hides, and horn; copper, gold, silver from ores containing these diverse metals; cotton fiber and cotton seed. Commodities produced at joint cost present peculiar problems of price.

A perfect example of joint production is that of cotton fiber and cotton seed. To make the fiber marketable the seed must be separated from it; all the expenses of cultivation and of ginning are necessarily incurred for the two together. But the prices per pound at which fiber and seed sell are very different. For every pound of lint (fiber) there are about two pounds of seed. At the prices of a five-year period (1903–08) the fiber sold at about ten cents a pound, the seed at about one half cent a pound. It may be assumed, with little divergence from the facts, that cotton is produced under conditions of competition, and that there is a wide margin at which the cost is practically constant. Fiber and seed between them therefore sell, taking their average prices over a series of years, for what it costs to produce them. But the apportionment of this total price between the two joint products depends on the relative demand for them or, in the terms which we have learned to use, on their marginal vendibility. The marginal vendibility of the cotton fiber from a given crop is much greater than the marginal vendibility of the seed produced along with it; hence cotton sells at a much higher price per pound.

It follows that an increase of demand for a commodity which is produced jointly with another may cause a fall in the price of that other. If the demand for cotton fiber increases its price will rise. This will not directly affect the price of seed, for which the supply and the conditions of demand remain the same. But the higher price of cotton is likely to stimulate production, and more both of fiber and of seed will be brought to market. The conditions of demand remaining unchanged for seed, its price must fall as supply is enlarged. Production will be increased until in the end the two between them will again sell for their joint expenses of production. But as the seed now sells at a lower price the fiber must sell at a somewhat higher price, and the definitive outcome of the greater demand for fiber will thus be a larger output of both constituents. It will cause a higher price for the one and yet entail a lower price for the other. The opposite effect would follow if demand for one of the articles should become not greater but less.

In most instances of joint cost, the situation is not so simple as this, for usually each article entails some separate items of expense. It is rare that, as with cotton fiber and cotton seed, all the expenses are incurred to the very last stage jointly for the two. The common case is more like that of wool and mutton; tho produced in the main at joint cost, each brings some special expenses of its own. The wool must be sheared, the sheep must be slaughtered and dressed for mutton. Wool and meat must each sell for at least the special cost connected with them, and for each a minimum price is thus set. In what proportion the remaining (joint) cost will be secured from the two will then depend on the play of demand, as in the simpler case of cotton fiber and seed.

The phrase "by-products" is often applied to denote some of the commodities produced at joint cost. When one of them habitually sells at a much lower price than the other, it is spoken of as a by-product; or when a material for which no use has been known comes to be utilized and to have a market value, it is so described. Both reasons explain why cotton seed is commonly spoken of as a by-product and not, as in strictness it should be, as a joint product. One of the most striking instances of joint cost is in the utilization of the various parts of slaughtered animals. The hide,

the bristles, the bones, the horns, the hoofs, the blood, the various organs, all are turned to some sort of use—usually with items of special cost pertaining to each. As the meat is the most important and familiar product, the others are commonly called by-products.

The advance in the arts of production, especially under the influence of chemical science, has led to the utilization of many materials previously thrown away, and so has made the principle of joint cost of wider and wider application. Wool, produced at joint cost with mutton, further illustrates also this aspect of the principle. As wool comes from the sheep's back it contains much fatty matter which must be got rid of before the fiber can be used for textile purposes. This matter, formerly waste, has in recent times been extracted, in some degree refined, and "lañolin" has proved an excellent unguent for several purposes. Similarly, cotton seed, itself a joint product, supplies not only the oil pressed out of it (and that oil of various grades, serviceable for various purposes) but also the oil cake remaining after extraction, which is used as food for cattle. The slag which comes to the surface of the molten matter in a pig-iron furnace, and of which vast quantities formerly accumulated near the furnaces (some parts being perhaps turned to account locally as ballast under railway ties), has come to be widely used as a material in the manufacture of cement.[1] Coal tar, one of the by-products from the making of gas and coke, has been found by chemistry to contain the materials for cheap and effective dyestuffs and also for important drugs. The crude oil which comes from the coal-bearing strata, and which has formed so wonderful an addition to man's resources during the last half century, is the basis of a number of products having partly joint cost and partly special cost—kerosene (illuminating oil), naphtha, gasoline, lubricating oil, dyes, paraffin and candles, vaseline.

For the utilization of some joint products a large plant is indispensable, as in the case of wool grease or coal oil products. In so far the advance of the arts has promoted the growth of large-scale production, and so has intensified the social problems which

[1] In Germany the slag left by the Thomas and other basic processes is the most important source of supply of phosphorus used as fertilizer.

arise from it. Large-scale production, in turn, may lead to monopoly or largely facilitate it. Then another complication appears. Either monopoly alone or joint cost alone entails consequences for value which diverge far from the simpler cases. When the two are combined, a variety of interacting forces must be considered—joint and separate cost, marginal vendibility and elasticity of demand, monopoly and maximum profit, and the effects upon monopoly of possible competition, of public opinion and public regulation, and of inert management. The Standard Oil Company in the United States illustrates all these complications. It long had a more or less effective monopoly, due to various causes, among which large-scale production and the utilization of joint products played their part; and these various joint products were marketed at prices influenced by all the factors mentioned in our discussion of monopoly, except probably that of inert management.

§ 2. A different case from joint cost is joint demand, where what is wanted is not a single article but a combination of articles. Thus a demand for dwellings is a demand for the completed accommodation. The purchaser is indifferent to the prices for brick, wood, glass, hardware; all he looks for is the house which combines these various materials.

If we suppose an increase in the demand for houses in a given district, and a rise in their prices, the change will be reflected in a rise in the prices of the several materials. If the materials were used solely for the construction of houses, and if they were put on the market under the same conditions—all equally limited in supply, or all equally extensible in supply—there would be no reason for expecting a greater rise in price for one than for the others. But the conditions of supply, as of demand, are likely to be different for the several constituents. Some may be obtainable in unlimited quantities with ease and at short notice; some may be temporarily or permanently limited. So far as any constituent is solely devoted to the given purpose and is limited in supply, so far is it likely to be peculiarly affected by the changes in demand for the joint product. Those constituents which serve other purposes also, and hence are on the market for miscellaneous sale,

will be diverted toward the joint product by the increase in price; enlarging supply here will check in some degree the rise in price. If the supply of any constituent be unlimited and easily extensible at constant cost, its price will not rise at all. Supply will promptly respond to the new demand, and the effect of that demand will appear solely with the other constituents. And if all the constituents except one be easily procured in larger quantities, and if their supplies thus respond quickly to an increased demand, that exceptional constituent will get the full benefit of the increase in price.

The different kinds of labor needed in building operations, as well as the different kinds of materials, illustrate the working of joint demand. A demand for houses and business premises means a demand for all kinds of workmen—for unskilled laborers, for bricklayers, masons and carpenters, for plumbers and electricians, and (in the case of high structures in American cities) for ironworkers. Some of these occupations are so widespread that an increased demand for a particular kind of labor in any one place easily draws an increased supply. This is most obviously true of ordinary manual labor—plain pick and shovel work. More of it can usually be got with little difficulty from other places. With the rougher kinds of carpenter's work the situation is similar. But it is different with the highly skilled trades and with those to which access is fettered by trade-union restriction. Here it is more difficult to add to the labor supply. Hence increased activity in building may have the effect of very greatly raising the wages of the workmen in these groups, while bringing comparatively little change for the others. Such a result has in recent years appeared frequently in American cities, strikingly so in New York. The rapid growth in urban population, combined with great improvements in building methods, has brought about astonishing activity in adding to and in remodeling dwellings and business premises. Certain kinds of laborers, not easily increased in numbers by recruiting from other occupations or from other places, have been in insistent demand—such as plumbers, tile workers, electrical workers, housesmiths (i.e. structural ironworkers). These have felt more than the others the demand for the joint product and

have secured extraordinarily high wages. Artificial restriction of the supply by trade-union regulation has sometimes played no small part in securing for them an exceptionally larger share of the possible gain.

Ordinarily, joint demand has not the same sort of permanent effect on value that joint supply has. In the long run, the conditions of supply are the more important in affecting value. Tho it is true, as appears most strikingly in the cases of increasing cost and monopoly value, that there is a constant interaction of supply and demand, the ultimate dominant forces for most commodities are those of supply. Where an increased joint demand affects most strongly some one commodity or some one kind of labor, because that happens to be the constituent whose supply is least easily extensible, there is none the less likely to be an increase in its supply. A readjustment of value takes place of the same sort as would have taken place if the demand had been not joint, but solely and separately for this one thing. If more brick is wanted more will be produced; and an increased demand for houses, tho it may for the moment raise the price of brick, will not do so permanently. But the situation is different with joint cost; an increase in the demand for cotton fiber may have a permanent effect in lowering the price of cotton seed. The immediate effect of an increase of demand is usually greater in case of joint demand, but the ultimate effect is usually greater in case of joint supply.

CHAPTER 16

THE INDIVIDUAL FIRM[1]

§ 1. The individual firm in the recent development of theoretical analysis. The firm's management of production here considered, not trading operations.—§ 2. The generalized principle of diminishing productivity, or law of varying proportions. Land, and seed applied to land.—§ 3. Other factors than seed, other industries than agriculture.—§ 4. The short-run cost curve of the individual firm.—§ 5. The short-run adjustment of output to maximization of profit. The long-run adjustment.—§ 6. The position of the individual firm different according as it makes a standardized or a differentiated product. Marginal cost and marginal revenue.—§ 7. Conclusion. Bearing and importance of this type of analysis.

§ 1. SOME attention must now be given to the economics of the individual firm, and to the relation between the operation of a given firm[2] and the economic structure as a whole. On this phase of economic theory great changes in the forms and ways of analysis have been proposed since the first edition of the present book was prepared. There are aspects of it which seem to me to be still in need of further analysis and elucidation. Some readers may wish to omit the present chapter, reserving consideration of its subject to a later time when they may take it up in connection with a more intensive study of these new developments of economic analysis.

In a system of private property and free enterprise the deliberate planning of economic activity is left mainly to the individual producers or firms. The manager of each firm can increase its profits in two ways,—more successful trading and more efficient production. Attention will be confined in this chapter to the operations of production. The trading operations give rise to problems somewhat different,—such as those of advertising, of unbridled competition, of monopoly and attempts at monopoly. Concerning these it is often difficult to say how far the business

[1] For the draft of this chapter, essentially as it stands, I am indebted to Dr. John M. Cassels; and he as well as I have to acknowledge heavy indebtedness to the treatment of the topic in Professor John D. Black's *Production Organization*.

[2] *Firm* is used here, and thruout the chapter, to mean any business unit, whether it be a single person, a partnership, or a corporation.

manager, in endeavoring to increase his gains, adds to the general welfare. But the endeavor of each firm to maximize the effectiveness of its own producing outfit may be said to lead to maximization also of the efficiency of the economic structure as a whole. The problem is in this respect a comparatively simple one; but it is by no means so as regards the *ways* in which efficiency is increased.

To the business man making decisions with respect to the management of his own particular enterprise, the problems with which he is confronted naturally appear to be unique. No other firm is faced with precisely the same set of conditions; no other has exactly the same set of personal factors to deal with: the same location, the same plant, the same market connections, the same industrial relations, the same financial arrangements. Yet there are enough essential characteristics common to the individual cases to afford a basis for a generalized form of analysis and the statement of some fundamental principles.

§ 2. Most fundamental of all is the principle of diminishing productivity for any one factor of production; or, as it is stated with greater precision, the law of variable proportions. It is in the nature of a physical law relating to the quantities of product obtained when, with given techniques of production, varying proportions of the different elements of production are used in combination with one another. The operation of the principle was first explicitly recognized in economic literature in connection with the diminishing returns obtained from the successive doses of labor and capital applied to land as the intensity of cultivation increased. Its application has been extended to cover all cases where the proportions of the factors of production used are capable of variation.

Suppose, for example, that seed is to be taken as the variable factor, while the area of land, the amount of water, the amount of fertilizer and the labor applied in cultivation are taken as fixed. Experiments have shown that, as the amount of seed is increased from small beginnings, the total amount of product obtained will increase for a time at a rate even greater than that at which the quantity of seed is increased; that then there is a stage in which

it will increase for some time at a rate less than proportionate to the increase in the seed; and that, finally, a stage is reached in which the total output of the product would actually be decreased by any further increase in the amount of seed that was sown. These may be referred to as the three stages in the operation of the law of variable proportions. It is only within the second stage that rational combinations or proportions are found. It would be irrational either to continue the increase of the application of the variable factor into the third phase or to stop short with proportions which belong to the first stage. Even if seed were available in unlimited quantities, free of charge, it would clearly be a mistake to use more than that amount with which the maximum total output would be obtained. If on the other hand the quantity of seed were rigidly fixed, and were insufficient when spread over all the land to give that intensity of application which marks the end of the first stage and the beginning of the second, a greater total crop would be obtained by concentrating the applications on a part only of the area, the other part of the land being left entirely idle. While farmers and other producers at times make serious mistakes in the factor combinations they employ, on the whole the actual combinations fall within the second stage of the law. It is because this is in practice the important stage, and because thruout its range the additional (or marginal) outputs obtained from successive increases in the quantity of the variable factor used show a decline of marginal output as well as of average output, that the law is commonly referred to as the principle of diminishing productivity.

The relations characteristically existing between the inputs of any variable factor and the outputs of product obtained thruout the second phase (and relevant parts of the other two) can be

TABLE I

Bushels of Seed	4	5	6	7	8	9	10	11	12	13
Bushels of Product	21	28	34	39	43	46	48	49	49	48
Average Product per Bushel of Seed	5.2	5.6	5.7	5.6	5.4	5.1	4.8	4.4	4.1	3.7
Marginal Product from Bushel of Seed	8	7	6	5	4	3	2	1	0	−1

best illustrated and explained by a simplified numerical example, and then summarized conveniently in graphic form by drawing three curves based on the figures assumed. Such figures are given in Table I and the corresponding curves are shown in Figure 10. The area of land on which the seed is sown may be taken to be one acre. As the amount of seed sown is increased up to six bushels,

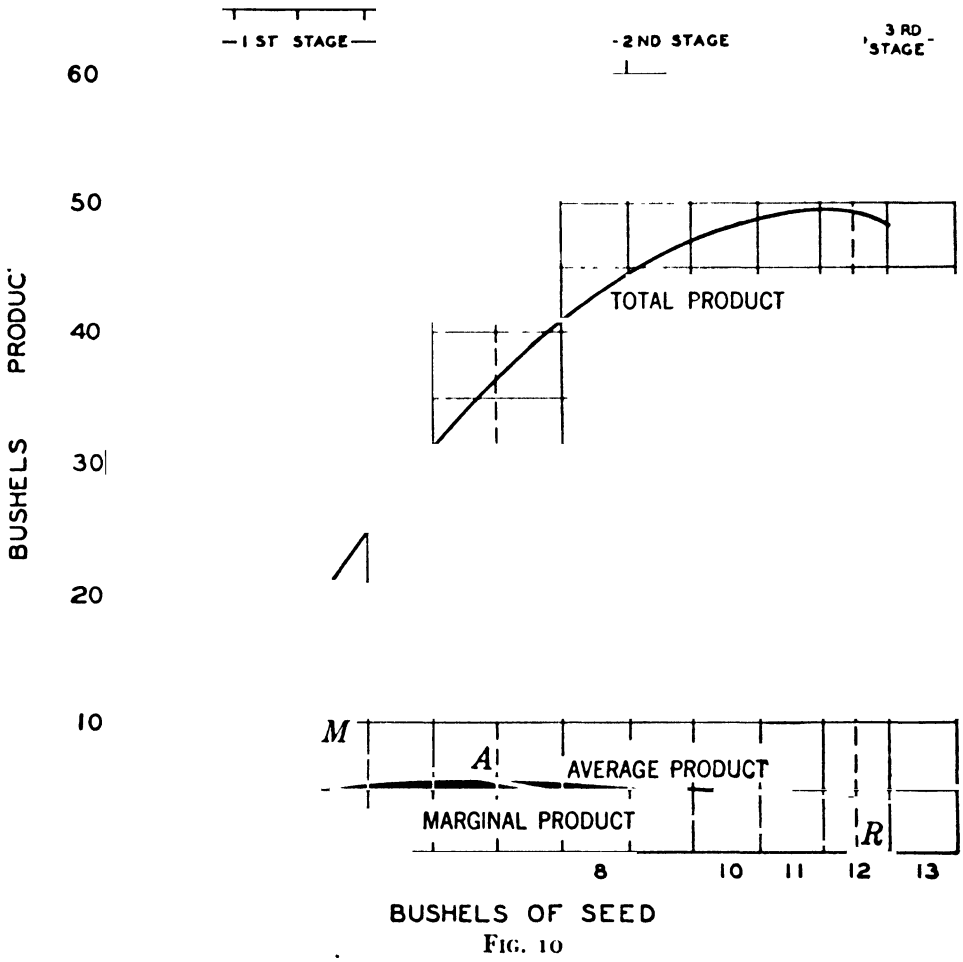

FIG. 10

the total output at each successive step is increased more than proportionately. Here the first stage ends and the second one begins. Further increases in the amount of seed continued to increase the total output until after the twelfth bushel has been applied; but thruout this second phase the *rate* of increase in output falls progressively farther below the rate of increase in the inputs of seed. With the twelfth bushel the rate of increase becomes zero. This last input of seed adds nothing to the output, the

marginal product is nil, and thereafter (if the procedure were continued into the third phase) it would be negative. The point of intersection of the marginal and average output curves (*A*) marks the beginning of the second stage; and the end of that stage is marked by the point (*R*) where the marginal outputs become zero. It is at (*A*) that the average curve begins to decline. The highest point on the marginal output curve (*M*) is reached before the end of the first stage. In the first stage, marginal outputs are greater than average, while in the second stage these relations are reversed. Both average and marginal returns decrease all thru the second stage. Under the conditions of production represented by these figures, there could never be any advantage in sowing more than eleven bushels of seed per acre or less than six.

§ 3. To determine exactly the amount of seed the use of which would yield the maximum profit (as distinguished from the maximum physical output) it would be necessary to know, in addition to the facts already indicated, the cost of the seed and the price at which the product could be sold. Not only that; in order to calculate the profit it would be necessary to know also the costs of all the other factors used in the production of the crop, such as land, labor, fertilizer, machinery. In other words, while the basic law with which we are concerned is purely physical in character, it must be combined for practical applications with all the relevant economic data, and thus made significant as an economic law of the relation between cost and price.

It is important to observe that in the foregoing illustration the assumption that the quantity of only one factor varies while the others remain fixed was merely an analytical device adopted as a means of simplifying the exposition, and that the choice of seed as the one variable factor is entirely arbitrary. Had labor, or fertilizer, or irrigation water, or any other element used in production been taken instead of seed, the general character of the results would have been the same, representing the operation of the same fundamental principle. In actual practice, the farmer must commonly consider the possible effects of simultaneous variations in a considerable number of the different factors entering

into his production combinations. He may use a tractor instead of horses and by plowing more deeply increase the returns obtainable from a heavier application of seed. He may use more fertilizer and thereby make it profitable to concentrate his own labor on the more intensive cultivation of a smaller area of land. He may substitute a combine harvester for an ordinary binder and thus economize on the amount of harvest labor required. Or he may change the combination of factors in various other ways.

Industrial producers are faced with similar problems. If they use more power, they may find it advantageous to use less labor; or they may discover that the greatest gain will come from maintaining the same labor force and increasing the supplies of raw materials bought for the factory. They must often choose between different types of machines, different kinds of labor, different grades of materials. The choice of techniques is usually greater for the industrial producer than for the agricultural producer, yet the basic principles involved are essentially the same. If we assume that the technical decisions are wisely made as the choices arise, and accept definitions of the factors of production which are not too narrowly restricted, it will be found that the principle outlined above has wide applicability and can carry us a long way in our study of the economics of the individual firm. It should be added that where several factors are varied simultaneously certain complicating intereffects are introduced, but for the purposes of the present discussion the explanation given above will suffice.

The nature and extent of the variations which a producer can make in the combination of factors he employs will naturally depend on the amount of time which is allowed for the adjustments to be made. If we think of a long period, one long enough for plant and machinery to be steadily used to the full until worn out and scrapped, we could also think of each firm or concern as working steadily with that combination of factors which is best under the technological conditions of the time.[1] But if we think of a

[1] We would assume also that the machines are marshalled; i.e., not all beginning together and discarded together, but with overlapping terms; the equipment of the industry as a whole remaining the same in kind, tho the output will vary as old firms gradually drop out more or less fast or new ones come in more or less fast.

shorter period--shorter, yet in many cases one of years and short only by comparison with periods of many years—it is clear that the plant and machinery cannot readily be changed. Durable factors such as these are for that "short" period fixed, not variable. Whichever these factors happen to be in any particular case, in any analysis of production costs they will occupy (for that period) a position analogous to that of the factors held fixed by assumption in our example. The other factors, those capable of being changed with comparative speed, taken together correspond to the factor which we have selected as the variable one, seed. Suppose the plant and machinery of a firm are fixed. A skeleton staff of workers with appropriate supplies of fuel and materials could no doubt produce some output of the finished product, but the operations would be carried on less efficiently than if the factory were more fully manned. As the inputs of these variable factors were increased the outputs per unit of input would increase until a certain point was reached and thereafter they would decrease. If the plant were overcrowded, night shifts were worked, and the pressure on the fixed factors increased excessively, the average returns would begin to decrease very rapidly. It is not likely that this pressure would be carried to the point where total outputs were actually diminishing, any more than it is likely that so much seed would be put in as to make the marginal crop output negative; but the possibility of this happening on occasions is by no means inconceivable.

§ 4. On the basis of this general reasoning about the changing efficiency of plant operations in physical terms, as the total output of the product is increased it is a comparatively simple matter to show what is likely to be the characteristic shape of the short-run cost curves for any individual firm. If we assume that we have data on the physical inputs and outputs, similar to those presented in Table I, then we need only apply to them the current valuations of the different factors in order to be able to work out for each volume of production the average overhead (or fixed) cost, the average variable cost, the average full cost, and the marginal cost. Hypothetical figures representing these cost relations for an extremely simplified case are given in Table II, and the correspond-

ing curves are shown in Figure 11, page 204. For simplicity it may be assumed that the firm pays the same price per unit for the factors, however large or small the quantity that it uses. Overhead costs, being by definition those which are independent of the volume of output, remain the same in their total amount. They

TABLE II

Output of Product	3	4	5	6	7	8	9	10
Average Overhead Costs	.83	.63	.50	.42	.36	.31	.28	.25
Average Variable Costs	4.17	3.62	3.34	3.20	3.15	3.19	3.27	3.42
Average Full Costs	5.00	4.25	3.84	3.62	3.51	3.50	3.55	3.67
Total Full Costs	15.00	17.00	19.20	21.70	24.60	28.00	32.00	36.70
Marginal Costs	3.00	2.00	2.20	2.50	2.90	3.40	4.00	4.70

include such things as interest on investment, depreciation due to weathering and obsolescence, taxes, and the wages or salaries of key men in the organization who must be kept on the staff if the firm is to continue in business at all. It is true that some of these items may sometimes be disregarded for years without cessation of operations; but sooner or later they must be faced, and they are included among the costs of production. As the output of the firm increases, the average overhead cost per unit of product tends to decrease. The curve representing these costs—they being for the short run a constant sum—must always be for the given period a rectangular hyperbola. As regards variable costs, their average per unit declines while the law of variable proportions is in the first stage of its operation, and reaches the lowest point (V) when an output of seven units has been produced. Thereafter they rise continuously.[1] Average full costs are obtained simply by adding together the average overhead costs and the average variable costs. This curve of average full costs is characteristically U-shaped.[1] Because of the fact that the overhead element decreases continuously, the output for which the average full costs are at a minimum (F) will always be greater than that for which the average variable costs are at their lowest. The marginal cost curve declines at first, but turns up (at M) before either the average variable cost curve or

[1] The curves are not extended to represent the operation of the law of variable proportions in its third phase. If they were they would bend backwards after the maximum total output had been attained.

the average full cost curve turns up and intersects both of them at their lowest points (V and F).[1] The same curve, it should be noted, represents the marginal full cost and the marginal variable cost. The marginal cost for any output might be determined in either of two ways: from the total full costs incurred in producing it might be subtracted the total full costs of producing the

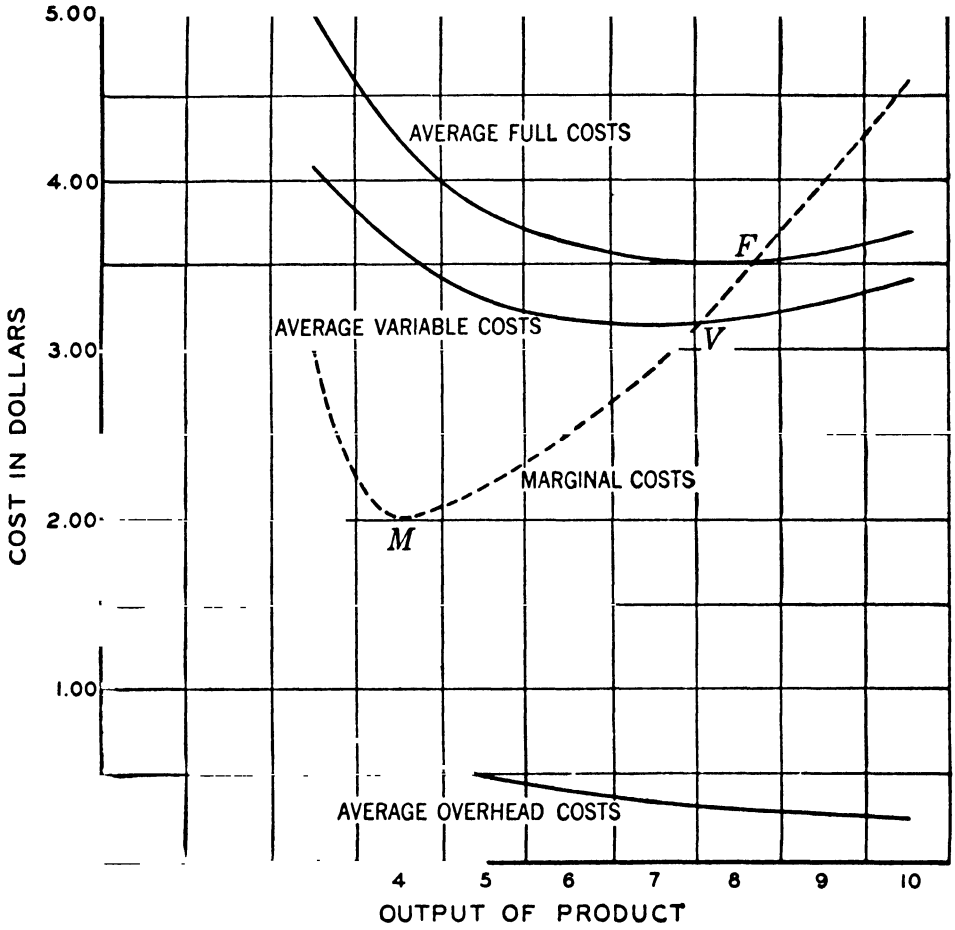

FIG. 11

output one unit smaller (as shown in Table II), or from the total variable costs of the given output might be subtracted the total variable costs for the smaller one. Since total overhead costs are always the same, the only element in total full costs which changes

[1] The slight inaccuracy of the figure arises from the construction of a continuous curve on the basis of the discontinuous data of Table II. The Figure is, I trust sufficiently accurate for the purpose in hand.

with a change in output is that made up of the variable costs. In other words marginal overhead costs are zero. When total costs $= K$ it follows mathematically that $mc = 0$. Therefore both of the above methods give exactly the same results and only one marginal cost curve is to be considered in our analysis.[1]

This technical and seemingly unrealistic terminology of economics is like that which has been invented by natural scientists for their distinctions and reasonings; having an exactness not to be found in everyday speech and thus conducive to clearness in exposition and discussion. In the world of actual business experience few men in charge of production, if any, express themselves in such terms. But these represent the general principles on which they do act. No doubt in most cases the individual producer has failed to be aware of them, still less to analyze them; yet by habituation he conforms to the principles which operate to determine the results obtained from the adjustments he does make.

§ 5. Suppose, now, that a firm were operating under the conditions represented by the hypothetical figures given in Table II, how should it adjust its output in order to obtain the greatest total amount of profit? To answer this question precisely and definitely, it is necessary to know, in addition to the information already given, the price or the prices at which the various quantities of the product can be sold. Here a distinction must be made between cases of pure competition and cases of monopolistic competition. In cases of pure competition the price obtained is unaffected (or only to a negligible extent affected) by the quantity put on the market by the individual producer in question; while

[1] The relations between average costs (ac) and marginal costs (mc) can be neatly set forth in mathematical form as follows:

if x = quantity, y = average cost, and e = elasticity of average cost curve, then since mc by definition equals $\frac{d(xy)}{dx}$, $mc = y + \frac{x dy}{dx} = ac - \frac{ac}{e}$

In the case of the average overhead cost curve the elasticity is unity and the formula shows the marginal cost to be zero (see text above)

$$mc = ac - \frac{ac}{1} = 0$$

For a fuller discussion of the mathematical relations involved, see Joan Robinson, *Economics of Imperfect Competition*, Chapter 2.

in cases of monopolistic competition the price he can secure will decline as he increases the supply of the product offered for sale. In the first type of case the analysis is naturally simpler than in the second.

We begin with the simpler type, taking a wheat farmer as our example. The wheat produced on his farm has no peculiar characteristics to distinguish it from the wheat produced on thousands of other farms. It must be marketed merely as part of a large homogeneous supply and cannot be sold at any higher price than that currently prevailing for wheat of that variety and grade. Since the contribution made to the market supply by any one individual producer is so small a fraction of the total, no change in his own output, taken by itself, can be supposed to influence appreciably the price that he will receive for it. A producer in this situation must simply estimate the price that may be expected to prevail when the crop is ready to be sold and make his plans accordingly (allowing as best he can for uncertainties in weather and so on). If the price is lower than the lowest average full costs attainable with any output ($3.50 with 8 units), there is clearly no possibility of making profits on the business of the year. If, however, the price is sufficient to cover the average *variable* costs, it will probably seem better, and indeed for a time will probably be in fact wiser, to produce some output than to suspend operation altogether. Should the price be lower, however, than the lowest average variable costs ($3.15 with 7 units of output) any production whatever would merely increase the losses. Assuming variable costs to be covered, the rule for ascertaining the output which would for the time being maximize profits (or minimize losses) is a comparatively simple one. The volume of output should be expanded to the point where every unit is produced for which the marginal costs are less than the price and none produced for which the marginal costs are greater than price. Suppose, for example, that the price is $4.25. Then it will be uneconomical to choose an output of 8 units, even tho that is the output for which average full costs are at a minimum, since the production of the 9th unit adds only $4.00 to the total costs incurred, and now brings in $4.25. On the other hand, to produce the 10th unit

would be unprofitable since the consequent increase in total costs is greater than the addition to the total revenue resulting from the sale of this unit ($4.70 against $4.25).

All this of course is not the whole story. The consideration of the short-run production adjustments of individual firms to maximize profits must be supplemented by analysis of the long-run adjustment. As time goes on, supply adjustments are brought about thru changes in the number and the size of firms engaged in producing different types of products or closely competing groups of products. If a firm steadily fails to cover its full costs of production, it will tend sooner or later to be driven out of business, thus reducing the total capacity in the industry. On the other hand, if any firm (or group of firms) is known to be making abnormally high profits, there will be a tendency for others to enter the field with additional products. Possibly there will also be a tendency for the earlier firm to enlarge its plant and expand the scale of its own operations. In these ways, productive resources are shifted over the long run from one field of production to another. The decline in demand for horse-drawn carriages turned profits to losses in the carriage-making industry, while the increase in demand for automobiles enabled successful car manufacturers to make enormous profits. The result in this case was not only a large withdrawal of resources from the one industry but a huge new investment in the other. The tendency for supply and demand to be brought into equilibrium by such shifts is beyond doubt one of the most important and fundamental characteristics of our existing economic system. It is not to be forgotten, or set aside as negligible, even tho short-run changes are probably becoming greater than in earlier days and are deservedly receiving more attention in the literature of economics.

§ 6. We turn now to the other case referred to at the beginning of this section where the relations between output produced and price obtained require special attention. For the whole of an industry, it is clear that an increase of output brings a lower price. But as regards an increase of output at the hands of the individual firm, this proposition needs to be stated in a somewhat different way, or rather with some qualifications. More specifically, it needs

to be restated if the firm produces not a standardized article of known quality and established attractiveness but something in the nature of a specialty.

Suppose a firm produces a commodity (or provides a service) which cannot be precisely duplicated by any other firm. It then sells in a more or less distinct sub-market of its own where the operation of the principle of diminishing vendibility applies directly and unmistakably to its own goods. The greater the degree of differentiation between its product and the nearest possible substitute, the more effectively will its own sub-market be cut off from the broader market in which all products of the same general class are sold. The differentiation between products may result from various causes; such as form or quality of the thing itself, to a brand or trade name attached to it, the location of the store in which it is sold, services that are given with it, personal relations existing between the buyer and the seller. Whatever the basis of the differentiation, wherever differentiation is present to an appreciable extent, each individual firm will find that the price obtained for its product is affected by the size of the output that is offered for sale. Each is then confronted with a demand curve for its own particular product similar in its essential characteristics to the general market demand curve discussed in Chapter 9. The principle of diminishing vendibility operates in the sub-market as it does in the general market. In order to induce buyers to take more of his product in preference to those with which it competes, the individual seller must either find a way of shifting the demand curve by some method of sales promotion (e.g. an advertising campaign) or else he must lower his price. For the present we will disregard efforts directed specifically towards demand creation, and consider only the second of these alternatives—reduction of price.

The producer of a differentiated product is not in the same position as a wheat farmer. The farmer, to repeat, finds it necessary to estimate only the market price at which he might expect to sell whatever quantity of the one article—wheat—he individually decides to grow. But the producer of a differentiated product must form some estimate of the different prices he might expect to re-

ceive with varying volumes of that special article. This is, of course, a problem in which many factors of great uncertainty are involved; it is one with which business men are obliged to deal as best they can. In some instances, especially where the changes contemplated are slight, the individual might be justified in assuming that his competitors' prices would be maintained at their established levels. But in what is probably the great majority of cases, he must face the fact that other producers may react to his adjustments with readjustments of their own. The demand curves used by economists are commonly drawn up on the assumption that the prices of all products other than the one under consideration remain constant. While this is the most useful type of curve to employ for some purposes, in explaining the production adjustments of a specializing individual firm a closer relation with reality is attained thru a demand curve which represents the expectations of the producer after he has made whatever allowances he thinks necessary for the reactions of his competitors. It is in the light of these general expectations that his policy will actually be determined. A demand curve drawn up on either basis will almost certainly slope downwards to the right; one drawn up to take account of expected changes in competitors' prices will probably slope more steeply than the other.

Suppose, then, that a firm producing under the conditions represented by the data in Table II has the demand expectations represented by the second line in Table III (page 210). As each successive unit of output is sold, the direct effect is to add to the total revenue an amount equal to the price corresponding to that quantity. But the indirect effect is to lower the price that might have been obtained for all the earlier units. The net addition to revenue resulting from the sale of any unit depends on the relative magnitudes of these losses as compared with the price at which the particular unit itself was sold. The simplest way of determining this net effect is to subtract from the total revenue obtained when the marginal unit is sold the total revenue obtained when one unit less was sold. The marginal revenue series obtained in this way is shown in the last line of the table. Applying the general rule given above for maximizing profits, the output should be

expanded to the point where every unit is produced for which the marginal costs are less than the marginal revenue (which was

TABLE III

Total Output of Product	1	2	3	4	5	6	7	8	9	10	11	12
Marginal Costs	7.50	4.50	3.00	2.00	2.20	2.50	2.90	3.40	4.00	4.70	5.50	6.40
Expected Price	5.05	4.95	4.85	4.75	4.65	4.55	4.45	4.35	4.25	4.15	4.05	3.95
Total Revenue	5.05	9.90	14.55	19.00	23.25	27.30	31.15	34.80	38.25	41.50	44.55	47.40
Additional Revenue	5.05	4.85	4.65	4.45	4.25	4.05	3.85	3.65	3.45	3.25	3.05	2.85

simply the price in the earlier case of an undifferentiated product). No units are produced for which marginal costs are greater than marginal revenue. Then we find that the firm should now choose an output of 8 units. On the eighth unit, there is a net gain of 25 cents (marginal revenue of $3.65 against marginal cost of $3.40). To produce the 9th unit now would be unprofitable since, altho its price ($4.25) is still higher than its marginal cost ($4.00), the marginal revenue ($3.45) is less than the marginal cost.[1]

The analysis made in this chapter is important in that it illustrates the difference in the adjustment of production under two kinds of competition: first, "atomistic competition" (also termed "pure" competition and "perfect" competition), when the output of each individual producer constitutes only a small part of a homogeneous supply; and second, "monopolistic competition," when the product of the individual seller is sufficiently differentiated to give him a fairly distinct sub-market of his own. The position of a firm under conditions of "monopolistic competition" resembles that of a monopolist. The difference between them is one of degree (or definition) rather than of essential characteristics. In

[1] Theoretically, for the establishment of equilibrium in either the short run or the long run it is necessary that marginal cost and marginal revenue should be equal (mc = mr). In the short run it is sufficient if the marginal revenue is greater than the average variable costs. But for equilibrium in the long run, where the entry of new firms into the industry is unrestricted, it is necessary as a further condition that average revenue should be equal to average full costs (mc = mr and ac = ar).

neither case can another producer supply customers with exactly the same product, and yet in both cases there will be more or less effective competition from substitutes. No one except the Ford Motor Company can produce cars having that name, nor can anyone except the American Tobacco Company manufacture Lucky Strike cigarettes; yet we do not ordinarily refer to these as cases of monopoly. On the other hand, a gas company is commonly said to have a monopoly because it is given an exclusive

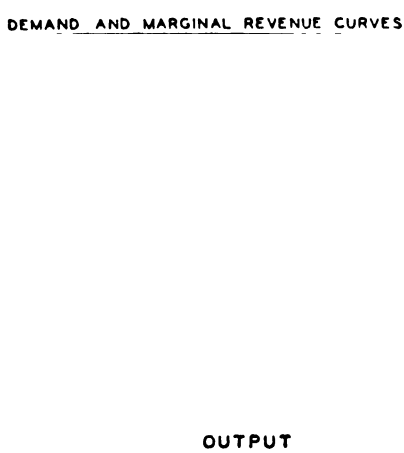

FIG. 12

right to pipe gas to houses in a certain district; yet this too is affected by the competition of coal, oil, and electricity. The distinction between those cases in which some element of monopoly is present and those in which the competition is atomistic may be simply stated in terms of the slope of the demand curve with which each individual firm is confronted. Under conditions of atomistic competition, the individual demand curve is horizontal (i.e. the price remains the same regardless of changes in the individual's output); whereas under conditions of monopoly or monopolistic competition, the individual demand curve has a negative inclination (i.e. the price declines as the output is increased). In the former case, marginal revenue is equal to price and the marginal revenue curve is identical with the demand curve, as shown in

Figure 12. In the latter case, marginal revenue is always less than price and the marginal revenue curve lies below the demand curve as in Figure 13. Bearing in mind what has been said—that

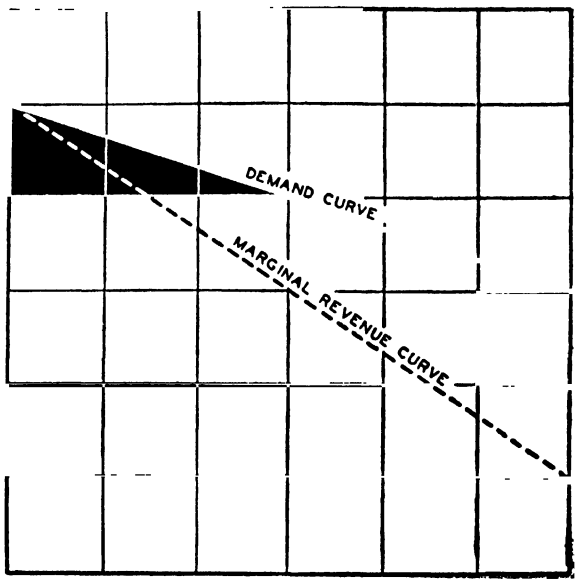

FIG. 13

it is not price but marginal revenue that directly determines the output which will maximize profits—the importance of the distinction will be readily perceived.

§ 7. In order that the most essential basic principles concerned with the economics of the individual firm may stand out as clearly as possible, the foregoing brief discussion has been made as simple as possible. Complicating factors have been ignored which would need to be taken into account in actual everyday experience. Considerations relating to production organization, personnel management, financial control, market strategy, competitive tactics and long-run planning are seldom simple in their character. Not even the assumption that individual producers will seek to maximize their profits can be accepted without qualification. There are firms whose ideal is long-run security rather than maximum profits at the moment; others in which an increase in the size of the business, the expansion of the industrial field under its control, comes to be regarded as an end in itself; and still others in which pride in the quality of service rendered or a conscious desire to con-

tribute to the general welfare are factors. Business behavior cannot be entirely explained in terms of cost curves and demand curves. Nevertheless these analytical tools go far toward promoting an understanding of the questions with which business men commonly have to deal. When they are closely analyzed and the qualifications needed in applying them are borne in mind, they can contribute effectively to the understanding of many economic problems of wide importance.

CHAPTER 17

MONOPOLY VALUE

§ 1. Monopoly affects price thru limitation of supply. This proposition qualified as to transactions between middlemen, especially as to producer's capital.—§ 2. How price is fixed if a monopolist has a fortuitous supply; how if he produces his supply at constant cost. Monopoly profit. Destruction of part of the supply possible, but not probable. Diamond mining as illustrating monopoly price.—§ 3. Monopoly price under increasing returns. Copyrighted books as illustrations. Monopoly price under diminishing returns.—§ 4. Possibility of varying prices under monopoly. Copyrighted books; telephone rates. Converse case of uniform prices under monopoly.—§ 5. "Dumping" explained by monopoly.—§ 6. Unqualified monopoly rare; various limitations and qualifications.—§ 7. Duopoly; complications of the theoretic problem; the complications even greater in the world of affairs. In what way and to what extent competition remains an active force.

§ 1. A monopolized commodity will be sold, by a person doing business for gain, on such terms as will yield the largest net revenue. We may assume, at the outset at least, that persons possessed of a monopoly act with shrewdness, and adjust their supply with intelligence and success so as to secure this maximum gain.

We say, adjust the supply; for this is the mode in which the monopolist can affect price and profit. The conditions of demand are beyond his control. When once the supply is settled and put on the market, the price at which it will sell depends on the play of demand. In this regard, monopoly value presents no peculiarities. Its special problems arise in so far as the monopolist can make the supply larger or smaller at will. With a given supply, put on the market *en bloc*,[1] the price will be the same whether it is in the hands of a single person or of several competing persons. There is some one price which measures its marginal vendibility—some one price at which the whole can be sold, and no more than the whole—and that price will rule.

This proposition, like so many in economics, needs to be taken broadly, as a statement of a tendency, not of literal detail; with precisely the same allowance for irregularity and imperfect adapta-

[1] See § 4 in this chapter for the significance of this qualification.

tion that must be made for any general statement on values and prices. If put as a statement not open to argument or qualification, most men in active business would deny it. They would say that a combination or monopoly can secure a higher price than competing persons can, even for the same supply. They know that a higher price can be obtained, in the first instance at least, from the middlemen, the wholesale or retail dealers, to whom the monopolist usually makes his direct sales. When producers are competing, these dealers are very apt to play off one against another and to induce the shaving of an offered price by threatening to turn to a competitor. No doubt, if all of the dealers do this successfully, competition among them will tend to lower prices in the end for the retail purchasers. At that final stage, it will appear whether the prices are such as to bring about the equation of supply and demand. But competition among dealers, and especially among retail dealers, operates with friction; and the lower prices which competition among manufacturers causes these to concede to dealers may redound for a considerable time to the dealers' profit, not to that of consumers. Conversely, a monopoly may squeeze the dealers, so to speak; charge them higher prices, which yet they do not find it feasible, for some time at least, to pass on to consumers. And even when such a rise in prices reaches consumers, the effect on their purchases is not immediate or automatic. If indeed the rise is great and the demand for the commodity is highly elastic, a reduction in purchases will be prompt. The monopolist will find almost at once that he cannot sell the same supply at higher prices. But if the rise in price is not great, people will very possibly continue to buy for some time what they have been in the habit of buying. Tho uneasy and irritated by the higher charge, they may yet for the moment not readjust themselves to the new situation by curtailing their purchases. The monopolist may then hold the raised price for a while, even tho consumers would eventually withdraw. Meanwhile, in a growing community, new consumers may be added, or the old consumers may get larger incomes. An increase in demand may overtake the higher price, and make it permanent; and then it will seem as if the mere fact of monopoly had caused prices to rise.

The position of middlemen as buffers, easing and delaying the pressure of the forces at work, appears even more strongly in the case of producer's goods. As has already been said,[1] the play of demand and utility is much modified in the prices of such things as iron, copper, timber, wool. The connection between the price ultimately paid for finished goods by consumers and the ruling price for materials among dealers is often a slow and uncertain one. Still slower is that between the materials for tools, like iron and copper, and the consumable articles which in the end the tools serve to make. Here there is a possible influence of monopoly on price which would not appear if the monopolist sold an enjoyable commodity directly to the consumers.

It is to be noted, further, that the first step taken by a monopolist is usually to settle his price, not his supply. The holder of a patent, for example, will offer the patented article at a given price; he will not usually determine in advance the amount which he will put on the market. If he finds that at the given price he can sell more than he expected, he will add to the supply. If he finds that he cannot sell so much, he will let the stock which he has on hand go off gradually, and in the future will add to it slowly and cautiously. In other words, he experiments with the supply which he can dispose of at the price fixed, and as time goes on lowers or raises his price according to the response from purchasers. Probably he is only half conscious that his control over price rests on his control over supply. Yet the shrewd business man is rarely in doubt that this is the fundamental condition for keeping a price above the competitive level.

§ 2. The power of a monopolist over price being exercised fundamentally thru his control over supply, let us examine further in what way the control is exercised.

The simplest case is that of a supply which has cost nothing—something in the nature of treasure trove. Such a fixed supply, if put on the market as a whole, will fetch a given price. But the owner may reason that a less supply will fetch a higher price. If the demand be inelastic, half of the supply may fetch more than double the price, and so yield a larger gross sum. It will then be in

[1] See above, Chapter 10, § 5.

the interest of the monopolist to destroy half the supply, and put on the market only the remaining half. If the demand is elastic, it will more probably be to his advantage to put the whole on the market. The price per unit, to be sure, will be lower than if only half were sold, but not so much lower as to make the gross yield less. It is usually to the interest of a monopolist to restrict sensibly the supply of a commodity subject to inelastic demand, and to be liberal with the supply of one subject to elastic demand.

Suppose next that the supply is not fortuitous, but is produced by the monopolist under the ordinary conditions, with capital invested, laborers hired, sundry expenses of production incurred. Then the monopolist will aim to obtain not the largest gross amount but the largest net profit. And that net profit he will try to make larger than the usual profits of capitalists. It may be assumed that in any case the monopolist would be able to secure on his capital, by investment in other directions, interest at the usual rate; and that for his own labor of direction and superintendence he would be able to secure the reward usually accruing to labor of the same skill and assiduity. Those normal gains we reckon among the expenses of production, or at least not as due to monopoly. It is the excess above them that constitutes *monopoly* profit.

It is probable that few monopolists consciously separate their gains in this way. They rarely distinguish between monopoly profits proper and ordinary returns for their capital and labor. They simply rejoice that they pay dividends at twenty or thirty per cent, or are able to be munificent on salaries to themselves and their associates. If closely questioned, however, they would soon distinguish the share in these gains which is due to monopoly alone. It is that share, monopoly profits in the strict sense, which now interests us.

If the monopolist produces his commodity under the conditions of constant cost, his calculation of net profit will be simple. Figure 14 will illustrate it. The cost of producing the commodity is there represented by the distance from O to C, and is the same whether a large or small amount of the commodity be produced; it is $OC = AC' = BC''$. The price at which any given quantity will sell depends on the conformation of the demand curve DD'. If a quan-

tity OA is put on the market, it can all be sold at the price AA'. The total cost of this quantity is $OCC'A$. Monopoly profit will then be indicated by the area $CPA'C'$. But if the quantity OB is put on the market, the price must be lowered to BB', that being the price at which the whole quantity OB can be disposed of. Monopoly profit is now the area $CP'B'C''$. If the first area, $CPA'C'$, is the larger of the two, it will be to the interest of the monopolist to restrict his output to the quantity OA. But if the

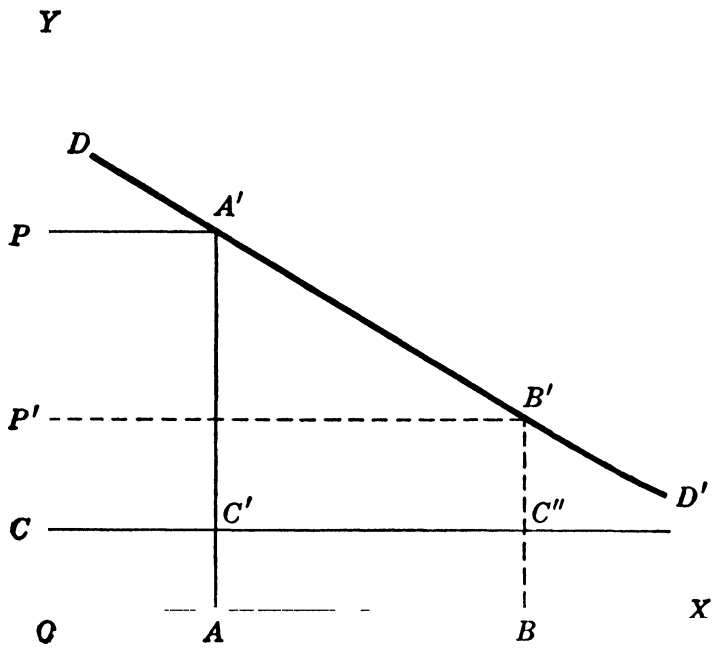

FIG. 14

area $CP'B'C''$ is the larger, it will be to his interest to enlarge his output to the amount OB. As has already been said, the elasticity of demand has an important influence on the calculations of the monopolist. If demand is elastic—if a lowering of price will greatly stimulate consumption and purchases—the quantity which can be disposed of at the price OP' will be greater than OB. The parallelograms indicating gross receipts and monopoly profit will be longer, and larger in area. Under such conditions it is probable that monopoly profit will be larger for a comparatively low price than for a high one.

In a preceding section it was said that a monopolist might find it to his advantage to destroy part of a supply, in order to sell the

remainder for a larger gross amount. But such destruction can take place very rarely. Fortuitous supplies, coming into a monopolist's hands without cost, hardly ever occur. When a monopolist's supply is produced and costs something, it is obviously easier and cheaper to refrain from producing a part of it than to destroy a part after it has been produced. Only from miscalculation or causes beyond control (such as superabundance of crops) may a monopolist find destruction to his advantage. It seems to be well established that in the eighteenth century the Dutch East India Company at times burnt part of its crop of cloves in order to be able to sell the remainder at prices so much higher as to increase its gross receipts. Similar destruction would hardly be ventured in a modern community; fear of retribution from an outraged public opinion would prevent it.[1]

The mode in which a monopolist commonly proceeds in the adjustment of supply is illustrated by the conditions of diamond production for many years. Virtually all new diamonds came from the mines at Kimberley in South Africa. These were under the single ownership of the De Beers Company, formed by an amalgamation, under the guidance of Cecil Rhodes, of a number of competing mines. Some of the mines were not worked, and the total supply was intentionally limited to the amount which could be sold to best advantage. The demand for diamonds, after a certain point, is highly inelastic. They are bought chiefly for purposes of display. Scarcity and high price are the basis of their utility; if very abundant, they would be little prized. Hence it was clearly to the advantage of the De Beers Company to curtail production and limit the supply.[2] Were the commodity one like copper, with a very elastic demand, it might pay such a monopolist to work the sources of supply to their utmost capacity.

[1] When a publisher prints a limited edition of a book, and then distributes the type, he may be said to wipe out part of the supply in order to sell at a higher price the copies which he prints.

[2] The De Beers Company controlled 95 per cent of the world's diamond production. See G. F. Williams, *The Diamond Mines of South Africa*, Vol. I, p. 291; Vol. II, p. 161.

Since the date of the first edition of this book (1910), the situation has changed somewhat. Discoveries elsewhere in South Africa have added to the supply, and the monopoly, while it remains effective, is in the hands not of a single producer but in those of a combination of a very few producers. The industry still illustrates the conditions of monopoly price.

§ 3. Suppose now that the monopolized commodity is produced not under the conditions of constant cost but under those of diminishing cost (increasing returns). The calculations of the monopolist then become complex. He must consider on the one hand the extent to which price will fall as a larger supply is put on the market, and on the other hand how much cost will fall as more is produced. The situation is again easily illustrated by a diagram.

On Figure 15 *DD'* represents an elastic demand. *SS'*, the supply curve, has a steep inclination, at least in its upper range, repre-

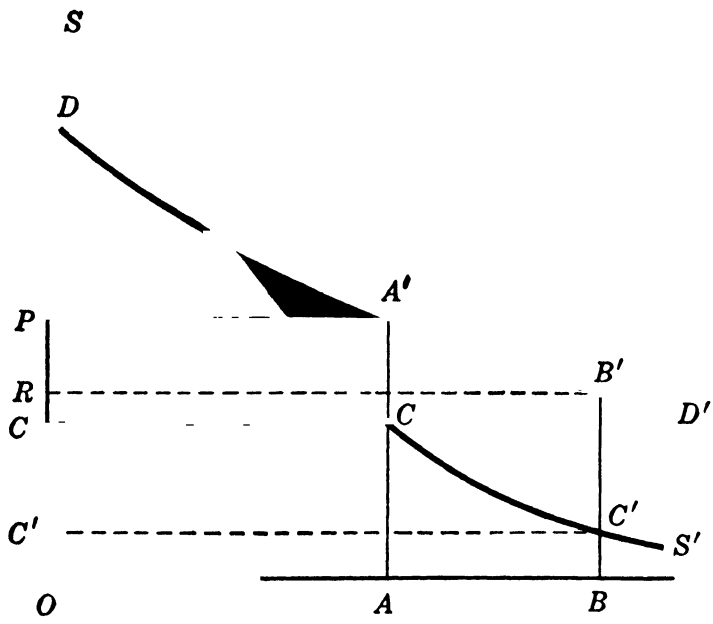

FIG. 15

senting a very rapid decline in cost per unit as supply is enlarged. If the monopolist produces and puts on the market the quantity *OA*, he will find the cost per unit to be *AC*, and the total cost to be *COAC*. That supply will be sold at the price *AA'*; the gross receipts will be *OPA'A*, and the monopoly profit will be *CPA'C*. If, on the other hand, the quantity produced is the larger amount *OB*, the cost per unit will be only *BC'*, and the cost of the total supply will be *C'OBC'*. That supply can be sold at the price *BB'*. The gross receipts will be *ORB'B*, and the monopoly profits will be *C'RB'C'*. Evidently the monopoly profit will be much greater with the lower price than with the higher price; this because the conditions assumed are those of very

elastic demand and of rapidly decreasing cost. The less elastic the demand, and the less rapid the decrease in cost, the more probable is it that the monopolist will find it to his advantage to limit the supply and keep up the price.

The reader will easily see that a number of maximum monopoly profits and ruling monopoly prices are possible. To express in one single statement all the elements of the case would require mathematical formulation. Such a formulation, however, has an appearance of accuracy which is often misleading; and this is true even of a comparatively simple diagram like that given above. Some of the elements in the situation must be more or less a matter of guess work for the monopolist; especially the degree of elasticity in demand, and the rate of decreasing cost with enlarged production. Even in the case of a perfectly unrestrained monopoly —and such are very rare—monopoly price is usually fixed by a sort of rule of thumb. Tho probably at a point considerably above the competitive price, it is not settled by any refined calculation of the precise point of maximum profit.

Sharply decreasing cost, or increasing return, is most likely to appear where articles are newly introduced. At first these are bought and used in small amounts. Later, as they become familiar and widely used, they are produced in larger quantities and the economies of large-scale production can be realized. Not infrequently new articles are monopolized, being protected by patent or copyright laws. They then give a most apt illustration of the working of the principles here under consideration. Thus, the Welsbach mantles attached to gas lights were long protected by patent in all advanced countries.[1] They enabled a much better light to be had for a less expenditure on gas, and they contaminated the air less. The demand for them was highly elastic. They were produced much more cheaply in large quantities. Hence, tho monopolized, they were sold at a price which, per unit of product, was not greatly above cost price; none the less, on the enormous quantity which could be sold, they yielded monopoly profits very great in the aggregate.

A situation essentially similar appears in the case of copyrighted

[1] This patent expired in the United States in 1906.

books. Any one book is produced more cheaply when large numbers are printed. The expense of typesetting and of making the electrotyped plates is the same whether one thousand copies be printed or fifty thousand. The other expenses of bookmaking—paper, presswork, binding, and the like—are tolerably uniform per unit, yet some of them also show diminishing cost as more copies are made, the cost per unit being much less for a large edition than for a small one. A common device of publishers is to issue a limited edition, often with numbered copies, and dispose of it at a high rate to collectors and other persons who prize the possession of a rare thing. They calculate that the profit will be greater from a small edition at a high price than from a large edition at a low price. The same result appears with scientific books, which often appeal to but a small circle of readers and for which the demand is inelastic. Only where books are salable in great numbers and the economies of large-scale production are thus attainable will the monopolist find it profitable to put many copies on the market and sell them at a low price.

Under conditions of increasing cost (diminishing returns) the situation of a monopolist will be different. The probability of a sharp limitation of supply is evidently greater if the increase of supply entails not less but greater cost for additional output. If the demand be highly inelastic, such a monopolist will certainly be disposed to restrict his output very much; for the price he can get will rise much with lessened supply, while his expenses per unit will fall. And even with an elastic demand he will have to reckon not indeed with rapidly falling price as output increases but with some increase in cost. Monopoly, however, with diminishing returns is probably rare. It may appear in the case of some uncommon mineral products, obtained from a single source of supply or a few combined sources (the South African diamond mines possibly presented an example). On the whole monopoly conditions, complete or partial, are much more likely to be found with commodities produced under constant or under increasing returns.

§ 4. Monopoly presents another possibility: different installments of the supply may be sold at varying prices. Under competition, one price prevails thruout the market; no one seller is

allowed by the others to get a higher price. In the preceding paragraphs it has been tacitly assumed that the same holds good under monopoly. But it does not necessarily hold.

Look, for example, at Figure 14 (p. 218) representing monopoly under the conditions of constant cost. The monopolist cannot but look with longing eyes at the possible profits represented by the area $CPA'C'$. It is true that the one uniform price yielding him the largest gain may be the price OP' $(= BB')$, at which his monopoly profits are $CP''B'C''$. But may he not get in addition the extra profit potentially to be had on the quantity OA, which would sell, if put on the market by itself, at the price AA'? May he not charge a high price to the richer or more eager buyers, while selling at a lower rate to those not able or willing to pay the high price?

To sell directly and openly at varying prices to different purchasers is, to be sure, not always feasible or politic. There is the possibility of resale by the favored purchaser. Moreover, the deep-rooted feeling for equality or "fair treatment" is to be reckoned with. Its violation arouses resentment which may affect purchasers or lead to hostile legislation. Hence the monopolist, if he discriminates, is likely to disguise his discrimination. In some degree he may secure from the upper strata of buyers that higher price which would otherwise inure to them as consumer's surplus.

The monopolist may put the commodity on the market in installments. He may sell at a high price first to those whose demand is keenest; and then, after a pause, put on the market a further supply at a lower price. Substantially this is often done by publishers with copyrighted books, especially such as are reasonably sure to have a considerable vogue. A first edition is offered at a comparatively high price. After a season or two, a much cheaper "popular" edition is put out, tempting a whole army of buyers for whom the first edition was too expensive. There is, indeed, some pretense of a difference between the two. The popular edition is printed on cheaper paper, has a less elaborate binding, may be in paper cover. But the difference in cost between the two forms is usually small and by no means accounts for the difference in selling price. That difference results in the main from the publisher's effort to tap in succession the several strata of buyers.

Something of the same sort happens not infrequently in the case of patented articles. These may be sold at a high price for the first installments put on the market, and at prices much reduced as the great mass of buyers are sought. There is, to be sure, another factor, already referred to. Being patented, the articles must be of a new sort; since the law gives the patent monopoly only on the ground of this novelty. The market is necessarily uncertain. The patentee is likely to proceed cautiously. Almost always it is only by gradual acquaintance that the public can be led to buy a novel article. The moderate quantity put on the market at the outset does not allow the advantages of large-scale production; hence, tho price is high, cost also is high. If it were certain from the start that the article would find a wide sale, large plant and elaborate division of labor might be applied from the beginning, great quantities might be produced, a small part sold at once at high prices, the rest stored away until it was time to satisfy the demand at lower prices. But this involves risk. Commonly, the earlier installments are produced and sold tentatively, and the advantages of low cost are not reaped until the possibility of large sales at low prices is proved by successive experiments.

A direct instance of discrimination in price seems to be supplied by the telephone. This is a monopoly in most communities, and indeed, whether under private or public management, ought to be a monopoly. The commodity, or service, is not of a transferable kind; hence one obstacle to discrimination—possible resale—is out of the way. Telephone rates are commonly adjusted on the basis of what the user can pay; they are higher in cities and in thickly settled districts than in rural districts. Some parts of the variations in charges are doubtless due to differences in cost, but in the main they seem to be the outcome of monopoly conditions. This form of price discrimination is not necessarily objectionable; it might be reached under public ownership as well as under private; it is adduced here simply by way of illustrating the peculiarities of monopoly prices.

A converse case occurs when a monopoly charges a level rate to all persons, under conditions which would lead competing producers to charge rates varying according to cost. Probably the

uniform fare in American street railways could not be maintained but for monopoly conditions. Custom, convenience in collection, and a disposition to conciliate the public account here for the one rate of fare which the monopoly charges. The most striking case of this sort is where a public authority carries on an industry as a monopoly. The uniform rate of postage on letters is to be explained largely in this way. The two-cent or three-cent rate is highly profitable on short-distance letters, and especially on letters in the large cities. If competing producers carried on the business, some of them would enter this profitable part of the field and carry letters there for much less than two cents. Other private individuals or corporations who might undertake letter service in outlying districts of thin population, especially the rural districts, would have to charge considerably more; or else the government would have to do the work at a heavy financial loss. The existing monopoly enables the government to cover the loss in one region from the profit in another. The postal service hence is administered at a very moderate uniform rate, usually with profit in European countries, at a comparatively small loss in the United States. The social and educational advantages of thus conducting the service as a monopoly with uniform rates are too obvious to need emphasis.

§ 5. The possibility of charging different prices to different purchasers explains the phenomenon of "dumping"—that is, the disposal of commodities in a foreign country at one price, and to domestic purchasers at another and higher price. In the absence of monopoly—that is, if producers were competing freely—all purchasers would get commodities at the same price. Competing producers might indeed gain collectively by selling part of the supply at a low rate and the rest at a higher; and where market conditions happen to be disadvantageous, and where the total supply cannot be sold on remunerative terms, there is a strong inducement to resort to such tactics. But no one producer will sacrifice himself for the benefit of the rest; he will not slaughter the whole or a part of his stock in order that others may gain. If however all were to carry out an agreement to sacrifice each a specified share of his supply, reserving the remainder for higher

prices, the object might conceivably be accomplished. Here, to be sure, there is this obstacle: a possibility that the favored purchaser may resell to those from whom it is proposed to exact the higher prices. But if the favored purchaser is a foreigner, and if a heavy duty on imports prevents him from sending back the "dumped" commodity to the domestic market, the obstacle is removed. The domestic price can then be kept higher, and the gain from this source may outweigh the loss on the dumped sales to foreigners; especially if the commodity be one for which the demand is inelastic and of which an increased supply on the domestic market would greatly depress the price. If the operation be carried on by a secure monopoly, it is possible that the foreign sales themselves will be at remunerative rates and that the higher domestic price will yield monopoly profits still further enhanced.

The more complete the monopoly, the more likely will be inequalities in the nature of "dumping." Even in cases of halfway monopoly or temporary monopoly, something of the sort may happen, tho the discriminations will be less striking and less continued. Any producer or vendor of a "specialty"—a particular brand, an unusual commodity—is apt to be for a time in a position of semi-monopoly. So far as he controls the given article, he may find it advantageous to get rid of part of his supply in a foreign country, or in any out-of-the-way region in order not to "spoil" his domestic market. Where control of the market rests only on good will, or on established plant and reputation, the extent to which dumping can be carried is obviously less than in the case of a firm and enduring monopoly. Where on the other hand many producers are steadily competing in the sale of a staple commodity, dumping will not arise at all.

§ 6. Complete and unqualified monopoly is rare. Hence too much stress should not be laid on the theory of monopoly price in explaining the phenomena of actual life.

A monopoly exercised by a government for fiscal reasons gives perhaps the best chance of exacting the full monopoly profit. When the Khedive of Egypt, in the days before the English occupation, maintained a monopoly of the salt trade, he probably squeezed out of it remorselessly all that could be exacted from his

unfortunate subjects. But fiscal monopolies do not generally exercise their power to the utmost. They are not uncommon in civilized countries, being simply a method of securing public revenue thru monopoly management instead of by taxes. Such are the tobacco and salt monopolies in the Austrian Reich, Italy, and Japan, the tobacco monopoly in France, the spirit monopoly in Switzerland. These are rarely exploited up to their maximum yield. A given net revenue, varying according to the financial needs of the several states, is sought and the adjustment of supply and of prices is pressed no further.

Patented and copyrighted articles, again, seem to fulfill the conditions of perfect monopoly; the law forbids competition once for all. But the holder of such a monopoly must reckon with the competition of more or less available substitutes, and thus is compelled to abate his prices and enlarge his supplies more than he would otherwise do. Copyrighted books for example must meet the competition of other copyrighted books of a similar kind, not to mention those on which the copyright has expired. A first-rate textbook yields a good monopoly profit, sometimes a very high one. Yet if the price be put too high, others little worse can be used in its place. The gain from a copyrighted or patented article often arises not so much from selling it at a higher price than others of a similar sort as from selling much more of it at about the same price.

In other cases, also, of real or apparent or half-way monopoly, there are commonly checks. Many so-called monopolies lack a legal basis and even a solid industrial basis. There is often only a weak basis with most of the "trusts" which have been formed by horizontal combination. They must be ever on the watch against competitors, and very few, if any, are in a position to exercise unrestrained monopoly power. Others, again, tho more securely established, must be on their guard against regulation or displacement by public authority. Such are the so-called "public service" industries—the railway, the street railway, the telegraph, the telephone, the gas companies. Both of these sorts of cases, so important in modern industry, will engage our attention as we proceed. Here it suffices to note that the monopoly is a qualified one.

Finally, the dullness or torpor of a monopolist must be reckoned with. The strict reasoning of the theory of monopoly price assumes him to press his advantage shrewdly and to the utmost. He may do nothing of the kind. The spur of competition—the one force which more than any other stimulates enterprise and business intelligence—is lacking. The secure monopolist is likely to be content with a good comfortable profit and to let well enough alone. It may happen, indeed, that another and shrewder person will see the possibilities, buy out the inert possessor, and proceed to manage the affair with more vigor and profit. Such has been not infrequently the course of events in the public service monopolies of modern times, especially those whose possibilities of profit have been connected with changes in the arts and with the rapid growth of great cities. But this is not a matter on which prediction can be ventured. The actual working of monopoly is often highly uncertain and irregular.

§ 7. What has been said about the difference between the close-reasoned theory of monopoly price and the uncertain working of monopoly in the actual world applies to a kindred set of phenomena, to which the term duopoly is applied. There may be more than one person (or firm or corporation) in control of the supply, yet not so many that each proceeds regardless of what the other does, as is the case with free or "pure" competition. Between them they control the supply but they do not act together as one man. Each has some regard to what the others put on the market. The simplest case and a representative one is that of duopoly, two producers and no more.

What would happen if this situation existed in all strictness—if there were (say) two eagle-eyed, cold-blooded, highly intelligent producers; no competitors; no regard to legislation or public opinion—has been much debated for a century by economists, especially by the mathematical economists. There is still no consensus of opinion among these whether the outcome of the calculations and doings of such duopolists would be a determinate or indeterminate price. The difficulty in the way of reaching a conclusion seems to rest on the circumstance that a necessary step in the reasoning is some assumption about the state of mind of the

two producers,—how each would expect the other to act. It is impossible in any particular case to foretell which among possible assumptions on this score is in accord with the facts, what is the way the duopolists will think and act.[1] Were this known, output and price would be determinate. Not being known, they are indeterminate. The output would probably be less and the price higher than under free competition, and the price probably lower than under monopoly. Conceivably it might rest at either extreme. There is no *a priori* ground for concluding just where it is likely to come between these limits. The difficulty becomes less if there be not two but several firms in control of the supply.

The sort of theoretical conclusion just stated holds if the duopoly is firmly established,—if the few producers have complete and unfettered control of the supply. The uncertainty which theory admits under these strictly defined conditions must be faced even more in the confused conditions of the actual world. Established and continuing duopoly is hardly to be found. True, there are cases where the evidence at first blush seems to show it to exist. But on closer examination there usually proves to be a cross between duopoly and horizontal combination. For example, the great steel concerns of the United States are few, headed by the United States Steel Corporation. Their relative quantitative positions in the industry remain about the same over considerable stretches of time. They have some formal understandings and doubtless further tacit understandings. But they watch each other and suspect each other. They act together, yet do not cease to struggle to hold each its own "fair share" of the business. They thus have earmarks both of duopoly and of a semi-monopolistic combination. Back of this there is another element: they are not sure how far they could go even if in complete agreement among themselves. They must still be prepared to face interlopers,—new, ambitious, able competitors. This kind of uneasy and unstable situation may persist indefinitely. Possibly it is only a stage of transition to an all-embracing monopoly, tempered by a few smallish independent

[1] The topic is too abstruse for elaboration in a book like this. For a critical history of its treatment by the mathematical economists, see the chapter on Duopoly and Oligopoly and Appendix A in Chamberlin, *Theory of Monopolistic Competition.*

concerns which are tolerated in order to veil the monopoly. A similar state of affairs would seem to exist in the great oil concerns. It does not fall readily into any one of the familiar categories.

The problem of duopoly arises with the greatest distinctness in certain manufacturing industries. The industries just mentioned —oil and steel—are highly integrated, there being vertical combination, based (in part at least) on control over natural sources of the supply. But there are many industries of the more common and representative sort, not highly integrated, where we find a small number of individual concerns, each on a large scale, each on guard against the others, all concerned chiefly with the working of competition among themselves and all having in mind some sort of sharing of the field. A cross-section of an industry thus constituted would show at any given stage—a year or a quinquennium—the earmarks of duopoly. Yet over a series of years, changes would almost always appear both as regards the relative positions of the old concerns and the parts played by new ones. The duopoly seems to be there for the time being; but it does not persist with the same make-up, and the changes in make-up may indicate a working of competition, delayed in its effects and imperfect yet always around the corner.

Underlying all that has been said in the preceding paragraphs is a point of cardinal importance on which the best-informed observer cannot speak with certainty. How far do the economies from large-scale production go? If these have their limits, the large concerns, tho few in number, do not constitute a circle closed once for all. According as the growth of the economies from size is limited or unlimited, there may be settled monopoly, settled duopoly, or an underlying competition, slow-working and imperfect yet in the end effective.

Not only these economic factors but political factors affect the situation. The further great-scale production goes and the closer the approach to monopoly, the more likely is public control, —thru regulation, taxation, outright ownership and management. Public control is in strictness an economic factor, just as the system of private property is. What deserves attention in the present con-

nection is that the mere possibility of public control is an economic factor, having its effect on the way in which private ownership and management run their course. What has been said on this score about complete monopoly is applicable to duopoly as well. Another complication is introduced, and the interpretation of the shifts and changes of modern industry, the ascertainment of what are anomalies and what persisting characteristics, become more and more difficult.

CHAPTER 18

SEMI-MONOPOLY. MONOPOLISTIC COMPETITION

§ 1. Brands, trade-marks, advertising. Their good side tolerably clear as regards producers' goods.—§ 2. Less so as regards consumers' goods. Costs made higher.—§ 3. Effects of the advertising of branded goods. Difficult problems in the strict field of economics.—§ 4. Even more perplexing problems as regards welfare and happiness. Questions as to the scope of economics.

§ 1. In the present chapter some problems of value will be considered which are on the border line between competition and monopoly and do not fit readily into any one of the categories of the chapters preceding. They relate to the effects of brands, trade-marks, labels, advertising.

Purchasers often look to some external mark in deciding which among comparable products they will select. They are especially likely to do so when separated from the makers by a long chain of intermediaries such as selling agents, wholesalers, retailers. They buy an article which has a brand or label or name which is supposed to guarantee quality. The producers find it necessary to create and maintain reputation and "good will." Often they find it profitable to do so and sometimes enormously profitable. Without some identifying mark or brand, large-scale production seems to be impossible. This obstacle in the way of developing a great enterprise does not appear with the standardized and absolutely impersonal raw materials, like wheat, cotton, pig iron, lumber, ores. These can be sold in huge volume anonymously, so to speak. Wheat needs no trade-mark, but in these modern times bread seems to. As goods become more and more specialized in the course of fashioning the raw materials, they acquire something in the nature of an individuality that must be made known and recognizable. The persons who produce them try to acquire a monopolistic or semi-monopolistic position.

Trade-marks and the like have their good sides and their bad; and the same applies to the advertising which is so conspicuous in

connection with them. In general the good sides appear most in the marketing of producers' goods and the bad sides most in that of consumers' goods.

First, producers' goods. They are of many kinds and many degrees of importance, ranging from simple mechanics' tools like saws, hammers, and screw drivers to large and intricate contrivances such as machine tools, spinning frames, automatic looms, diggers, dredgers. The persons who buy them and use them are usually well informed, know what they want and can also judge the quality of what is offered. If they are influenced by trade-marks, it is because of experience. A tool or machine or specialized material which does not prove as good as expected will not be bought again. To hold customers it is a condition of profit that good quality shall be maintained; the brand must be kept up.

In this as in all economic phenomena there are set-backs, eddies from the main current. Sometimes the maker of a well-approved machine gets into financial difficulties and acts with regard solely to the exigencies of the moment. He may be in debt to his bankers, perhaps so deeply that they have "taken over" his affairs. They may prescribe that the product must be made cheaper, even tho of poorer quality, and still put on the market under the same name and at the same price. The procedure may work for a while, and the bankers may "get out" thereby. But the trade-mark has become worth less, perhaps quite worthless. Every business man knows that this is short-sighted policy, and it is not often followed. The decline of an established brand among producers' goods usually sets in for other reasons, such as the invention of an entirely new and better device.

It is curious and also significant that even for established trade-marks of this kind some advertising is essential. Anyone who looks at a trade journal will see page on page giving names of firms and brands; sometimes the bare names, sometimes also mention of the age of the concern, sometimes elaboration of the superiority of the product. One might suppose that well-known articles of this type would be bought once for all when wanted and that egging on of purchasers was superfluous. But it seems not to be so. The purchasers must be reminded; the familiarity of a name, if only from

mere repetition, is an essential element in holding the market. There must be defensive advertising to meet the offensive of intruders.

On the whole, each producer tries to erect barriers to competition, and yet the process is a competitive one. What seem to be wastes of competition are the results of a tendency to lethargy among buyers, and are concomitants of progress: they keep both sellers and buyers alert. As regards producers' goods, the trade-mark introduces no new factor of importance into the theory of value or into the great social problems.

§ 2. In the case of consumers' goods, the phenomena are more complicated. Competition is more uneven in its working and more dubious as to its results. Here advertising plays an immensely larger part.

It is often said in business circles that it pays to advertise a good article. As regards producers' goods probably it does; the emphasis on goodness is justified. But as regards consumers' goods it often pays to advertise with little regard to the goodness of the article. In modern times the cheap newspaper and almost universal literacy have made it possible to reach thousands of people, even millions, who are ignorant and gullible. The radio has made it possible to reach the quite illiterate too. Blatant, mendacious, vulgar advertising confronts us at every turn.

Questions of two kinds arise: first as to the effects on costs, on prices, on competitive or non-competitive gains,—the questions with which we are here chiefly concerned; and second, questions much more difficult, as to the meaning of it all for the well-being of mankind.

As regards costs and prices, the situation for consumers' goods is in some ways similar to that which has been found for producers' goods. Competition is loaded, so to speak. Every purveyor to the consumers must have a distinguishing mark of some sort. It may be a trade-mark duly registered and protected by law from infringement; it may be a name or slogan; as regards a retail shop it may be merely a location familiar to customers. Each tries to have something that his competitors do not have, and so to get into a semi-monopolistic position. Hence the term "monopolistic compe-

tition" has been devised[1] as a label for the situation, indicating how there is an interaction between the two kinds or sets of conditions, and a transition thru many gradations from pure competition to complete monopoly. Everyone who caters directly to consumers finds himself compelled to make pretension of some special excellence and to spend money for the purpose. Without such expenditure, large retail operations are impossible. All incur the expenditure, and for each and every one the costs of doing business are jacked up. Some do it more effectively than others; and, as between business concerns, this is one among the differences which, taken together, are so important for the theory of business profits.[2] But all have to reckon the expense as an item in costs, and it must be recouped for all in the prices paid by the buyers.

§ 3. How far is this futile,—a mere forcing on each other of expenses which yield in the end no net gain to any one and yet enhance prices for consumers? The case is not so simple as it may seem. Sometimes a gain to the community does accrue from advertising. Advertising is essential for the introduction of things quite new, whose serviceability is not to be gainsaid but has still to be made known. Here a launching campaign, often costly, cannot be dispensed with. Another case, not quite so simple, yet not difficult to explain and accept, is the advertising so freely incurred for selling articles approved and familiar; for the maintenance of a tradition. It appears, as has just been remarked, for producers' goods as well as for consumers' goods. In the industries which are close to consumers it is peculiarly necessary to advertise in order to keep things going. Not only a launching campaign but a continuation of advertising seem to be inescapable adjuncts to business on a large scale. The incessant vaunting has a sort of indirect or preventive productiveness, analogous to that of the policemen and the criminal courts: directly they contribute nothing but without them things would not go smoothly and uninterruptedly. Apparently the advertising must be there in order that effective production shall be continuously maintained.

[1] By Professor Edward Chamberlin, in his notable book on *Monopolistic Competition*.

[2] See Chapter 19, Vol. II.

At the other extreme we find cases which also are simple; where there is palpable lying about the goods, acts clearly fraudulent, or the offering and lauding of articles clearly injurious to health or morals. These the law handles with comparative ease—logical ease —and with little hesitation. The problems they present to economics are not essentially different from those of larceny and burglary. There are rules of the game, laid down by the law of property, within which competition must be confined.

Midway are the puzzling cases, lying on the borderland where jurisprudence touches economics and the other social sciences. We find doings about which the law holds aloof; unwelcome, but not so plainly bad as to be ruled out once for all. Such are mere exaggerations, where it can be said with some composure that after all the purchaser must look out for himself. Most dubious are the cases where there is something other than mere puffing. The consumers' tastes and purchases may be not only aroused but be deliberately created: say by a suggestion of distinction conveyed. Shall it be said here also, *caveat emptor?* Here, further, there is competition in extravagant tooting. If the tooting is done more effectively by one than by another, it may establish and often does establish a differential position. A semi-monopoly arises, lasting for a period as long as the business community is usually concerned with,—lasting, that is, for the business generation. It is often associated with large-scale production; yet not necessarily so. It means high costs but may bring high gains. The whole performance is among the most repulsive of the characteristics of modern society. And it brings consequences in the prices of the lauded goods which are not easy to fit into the usual categories of economics. While launching expenses and expenses for maintenance of an established business trade-mark or connection can be fitted into the broad lines of the simple theory of value, here there are perplexities.

§ 4. The perplexities belong with the second class of questions referred to a moment ago, those which relate to welfare and happiness, to the ends of life. It is quite an open question whether these larger matters should be dealt with at all in a treatise on economics. The scope of economic science is admittedly limited.

The usual statement, the first approximation toward defining the science and the ends which it envisages, is that its concern is merely with what people want or desire. The commendability, the moral or esthetic goodness of their wants and satisfactions—all this is outside our scope. We examine merely the ways in which people produce and exchange and divide among themselves the things they happen to want. "Value" is pertinent only in the sense of exchange value, not in what is called moral or esthetic value. It is immaterial for economics how a want arises, or even whether purchases are made with the results expected or not; the only question is what people want at the time when the purchases are made and what will be paid. It is all a matter of vendibility and the relation of costs to vendibility, of disutility to utility.

It would seem to follow that the how and why of wants lies out side our field. This bringing of wants into existence, or at least into a quite new form of expression, is a queer thing. How to explain it stirs the scientific curiosity of psychologists and moral philosophers as well as of economists; and that curiosity itself is a queer thing! When the wants somehow are there, then the economic problems arise. The phenomena here under consideration have a special place in economics only so far as competition works slowly, imperfectly, unpredictably—so far as the relation between costs and prices is irregular and uncertain, semi-monopolistic conditions persist for a long time. There seems to be an accentuation of the trade-mark situation, but not a situation radically different.

What is to be said of the large gains and even great fortunes that sometimes arise? They seem not to be related to any corresponding contribution to welfare. But then we pause, and ask what is welfare? The factory girl buys rouge or cigarettes because they are vaunted to be the same as are used by the lady of fashion; that personage's highly paid and much-advertised certification may even prove beyond doubt that it is exactly the same. The girl likes it because it gives her a sense of distinction. Again, the man uneasy about his state of health buys a much-advertised concoction which is harmless, or a mild corrective well known to the medical profession. He may feel the better for using it just because he is

told to expect betterment. Blatant, vulgar, repulsive are the adjectives which I have just applied to some of the familiar ways of enticing consumers. But the terms after all are mere expressions of a personal or class preference, no nearer to universal validity than one's judgment on a book or a picture.

As was said before, it is the sudden development of almost universal literacy, and of newspapers and the like commanding almost universal attention, that has led to phenomena distressing to the educated and refined. They go with the breakdown of class distinctions and the sense of a new kind of freedom in social relations. It has often been said that the true goal of happiness is the utmost freedom, and the true test of a social order is its attainment of the utmost freedom. But what is that "utmost"? Some limits there must be. Is a freedom of choice such as brings the anomalous results we have been considering to be restricted? Are advertisers to be censored not only on matters of health but on matters of taste? Perhaps the unsavory things which of late have become so conspicuous are but the froth and scum of a rapidly changing society, likely to subside as it settles down to a new order based on more of education and intelligence and less of inequality. What is the wise way of sitting in judgment on them—the way that will commend itself to later generations?

We have thus been led to touch on topics that lie outside the usual scope of economics. They reach, to repeat, into psychology and into ethics and the search for the good; to reflections on what is happiness, what are the aims and ends of life. The economist's field, given the limits just indicated, is merely a part of the large one. While he cannot refrain from sometimes crossing the line which separates him from the other parts, yet in general probably he does best to keep within his own domain. Its problems are perplexing enough. They have puzzled acute thinkers, in many ways still puzzle them. They appeal to the curiosity of the scientific mind. And they immensely concern all men, the dull as well as the intelligent, the rich and the poor, the altruistic and the selfish. The first and strongest interest of men is in their physical needs; the next in their developing desires and the means of satisfying them. About these matters there are misunderstandings, errors, ill-

directed moves. The economist's judgment about them is not all, but it is not to be disregarded.

REFERENCES ON BOOK II

First to be mentioned on the theory of value is Marshall, *Principles of Economics*, Books III, IV, V (Edition of 1910 or of any later date). Also on a high level is Wicksell, *Lectures on Political Economy*, Vol. I, General Theory (English Translation from the Swedish, 1934), which incorporates re-formulations. The more recent literature on the general subject is voluminous but still (1939) mostly in the form of papers in periodicals. Among systematic restatements to be mentioned are: J. M. Clark, *Economics of Overhead Costs* (1923); J. Robinson, *Economics of Imperfect Competition* (1933); Edward Chamberlin, *Theory of Monopolistic Competition* (1933).

On the relation between value and welfare, see the admirable series of chapters in Pigou, *Economics of Welfare* (4th edition, 1932), Part I; another series in Part II, on monopoly conditions.

It is impossible to separate the problems of value from those of distribution, and the reader is asked to turn also to the References given in the second volume on the subject of Distribution.

BOOK III

MONEY AND BANKING

CHAPTER 19

INTRODUCTORY

§ 1. This Book takes up chiefly questions arising about the general level of prices. The value of money. What "money" means: coin, paper money, deposits.—§ 2. Coins and coinage. Gold and silver and their long use. Coins, formerly dominant, now in wide use for small payments only.—§ 3. Paper money and its use in modern times. Convertible and inconvertible paper.—§ 4. Bank deposits and the check system. The legal position of banks and of their depositors. Growth of deposits in English-speaking countries. Why classed as "money."—§ 5. The "quantity theory" of money. The demand for money has an elasticity of unity.—§ 6. Explanations and qualifications.—§ 7. The way in which money and goods meet each other in the market. Velocity of circulation. Custom the chief determinant in the circulation of money.—§ 8. It is not important whether the price level be high or low, but highly important whether it changes.—§ 9. How reckon the total quantity of money at a given time and place?

§ 1. OUTSTANDING among monetary problems is that of the general range of prices. As already explained,[1] and as indeed is obvious on a moment's reflection, the general range of prices is the value of money. A given sum of money buys many goods when prices are low; its value then is high. The same sum buys few goods when prices are high; its value then is low. The preceding Book has considered changes in the values of this or that commodity on the supposition that the general range of prices is constant. The present Book considers chiefly changes in the value of money itself; why and how it is that the general range of prices is not constant but moves up or down. While this is by no means the only topic of the chapters to follow, it is the central one. Monetary problems, both theoretical and practical, ramify in many directions, but all are connected primarily with the movements of the general price level and with the effects of stability and instability in that level.

Every one has a general idea of what money means: the thing wherewith he buys goods, pays debts, hires labor. It is most commonly thought of as that which passes from hand to hand for these

[1] See Chapter 9.

purposes, the coin and paper which we carry about in our pockets and purses. In ordinary talk we do not usually include under the term that which enables us to merely procure this "money," such as a debt due to us from another or a deposit in bank. And yet, as regards the bank deposit, we commonly think of it as "cash on hand." Strictly speaking, a bank deposit is merely a debt owed by the bank to the depositor. But it is a debt of a peculiar kind, being transferable by check to another person with such ease that it serves in most cases as well as coin or paper and in many cases better. The extraordinary growth of the use of deposits and the use of checks on them in making payment, and their enormous importance in the monetary mechanism of modern times, have led writers on economics to include deposits under the term "money." Deposits, be it noted, not checks; the checks are merely the deposits in act of use. On this topic much more will be said shortly. For the present it is enough to say that there are three kinds of money, or means of payment, in countries like England, the United States, Germany, France: not only the coin and paper money, but the deposits also. These three will be considered in the chapters that follow, each at some length. It is only after a detailed and careful consideration of their origin and working that anything like a solution can be reached on the main problem, that of the general level of prices.

§ 2. Coins are stamped and certified pieces of metal. They are manufactured and supplied almost universally by a public authority. In all advanced countries coinage is carried on as a public function. It might conceivably be left to private persons, as it is partly in China today; but this puts the users of coins to the trouble of having to judge of the weight and fineness of coins just as they do of spoons and forks. The convenience of coins as part of the medium of exchange is immensely promoted by uniformity, and uniformity is secured most easily when governments reserve to themselves an exclusive right of coinage, punishing as a crime the private manufacture of money pieces.[1]

[1] Historically, a strong reason for the public monopoly of coinage was the desire of kings and princes to make a profit by minting operations, often thru deliberate debasement. In modern times the monopoly is maintained because thru it alone uniformity in the circulating medium can be assured.

Further to secure uniformity and convenience, coins are so manufactured that they cannot be clipped or whittled without easy detection of the defect. Thus designs are put on both sides, and corrugations ("milling") or lettering on the edges. If the coins were simply round flat pieces of metal with smooth edges, shavings could be scraped or cut from them without easy detection. Such "sweating" was common in earlier days, before the art of coinage had been perfected. Modern machinery turns out pieces so skillfully manufactured that troubles of this sort have practically ceased.

That coins are made of metal is a matter partly of convenience and partly of economy. The fact that the metals can be split up into pieces absolutely uniform is one of the qualities which fits them for this use. True, the same quality is possessed not only by gold and silver, copper and nickel, but by other metals as well. The metals used at present have the further advantage of being subject in less degree than most others to rust and corrosion, while at the same time they are more malleable than such metals as steel. Gold and silver, the most widely used of all the metals for monetary purposes, have in their pure form one disadvantage: they are too soft to resist wear. Hence they are always strengthened by an alloy; a small percentage of something else—usually copper—is added to impart the necessary hardness and toughness. As actually used, the coinage metals have the qualities of divisibility and malleability necessary for careful manufacture, and at the same time sufficient durability to ensure long life and consequent economy in replacement. For money of small denominations, like the American cent and the dime and even for the quarter and the half-dollar, also the English shilling, the metals, because of their durability, can hardly be supplanted by paper, as they can be and have been for larger denominations.

Coins, once the dominant element in the circulating medium, no longer have any such position in the advanced nations. Their chief use today is in retail transactions. Here they continue to play an important part. They are convenient to carry about in purse or in pocket, are easy to recognize by sight or by touch, and are well adapted to machine-handling, as in theaters or street rail-

ways. It is improbable that coins will be superseded in this kind of use. But the volume of the dealings to which they are thus adapted is small in comparison to the whole bulk of transactions. Even in retail trade a considerable part of the sales involves the payment of amounts too large to be made conveniently in coins; here, too, paper has come to be preponderantly used as "cash"; and here, too, use of checks has grown apace. Coins have thus been reduced to a position of but little quantitative significance. In the United States they count up to only a very small percentage—usually less than 2 per cent—of the total circulating medium.

§ 3. Paper money in modern times has come to be used on an increasingly wide scale. It evolved from documents which came to be used for the transference with facility and safety of titles to stocks of coin or specie. In the first stages it was accepted in payment only because really exchangeable on demand for specie. Once people were habituated to seeing it received by everyone and paid out by everyone, it came to circulate with little reference to its convertibility into specie. In the world of the twentieth century, paper notes pass readily from hand to hand whether convertible or not. Their unquestioning acceptance may be promoted by one device or another, such as making them receivable in payment of taxes or other public dues, and a legal tender for debts; but it is not dependent on these. The ready circulation of paper money printed and sanctioned by public authority is now so imbedded in custom as to be a matter of course.

Paper money remains in outward form a promise to pay. That promise is by tradition and implication one to pay coin or specie. Sometimes the promise is not kept and the paper is not in fact redeemable in specie, in which case it is said to be inconvertible. Often enough, while the promise is not kept, there is an expectation that at some time in the future it will be. Where the paper money is really convertible there is held for its redemption a stock of specie.

As regards convertibility and non-convertibility there are extremes, ranging all the way from the most fully assured convertibility to complete inconvertibility. The conservative extreme appears in the cases (nowadays very rare and obsolescent) where

the same precise amount of specie is put aside and held intact as the amount of paper outstanding. This is the situation with the gold and silver certificates of the United States Treasury. Virtually this was done for the "non-fiduciary" notes of the Bank of England for nearly a century after the passage of the great Bank Act of 1844.[1] Here the only advantage of the paper arises from its convenience for the payment of moderately large amounts; there is no economy of gold from its use. More often convertibility has been assured by keeping a redemption fund of coin or specie, amounting to a moderate fraction of the paper, from which any holder can get the metal if he wants it. Thus in 1900 the United States set up a fund or "reserve" of $150,000,000 in gold for the conversion on demand into that metal of the United States notes, commonly called greenbacks, outstanding to the amount of roughly $346,000,000; an obvious economy in the use of the metal.

At the furthest extreme we have completely inconvertible paper, so-called fiat money, with no pretense of any payment or redemption in specie. Somewhat covertly, but more and more openly as time went on, inconvertible paper became familiar in the nineteenth century and common in the twentieth; sometimes issued and used in payments directly by governments, sometimes by public banks acting for governments. Here the economy is at the maximum, since the only expense is that of preparing and printing the paper money in such way as to give fair durability and to make counterfeiting so difficult as to become negligible. As regards general acceptability and readiness of circulation, there is in ordinary times no difference between the different kinds of paper. They pass from hand to hand equally well whether secured in whole or in part or not at all by the backing of a store of specie.

The modern world has become so accustomed to paper, and the custom has become so fully strengthened by the convenience of paper, that the use of coin seems to be relegated definitively for the small transactions where "change" must be made and durability is essential. The difficult questions concerning paper arise in other directions, not about its economy, convenience or ready circulation but about the amount of the paper which can be

[1] See below, Chapter 26.

wisely issued or created, and the possibilities of violent convulsions when it gets beyond bounds.

§ 4. The third constituent of the medium of exchange, and the most important in those places where banking is highly developed, is found in bank deposits. In the United States, all transactions of jobbers and wholesalers, all those of manufacturers except payments of wages, a large and growing volume of retail transactions, most salary payments and even a considerable part of wage payments, are handled by checks drawn against deposits in banks. All men of affairs, all members of professions, almost all salaried persons, all persons possessed of means, even a growing number of wage earners are users of banks. Much the same is true of England and of the British Dominions. Wherever English is spoken, bank deposits have come to form by far the largest part of the medium of exchange.

It may seem odd to speak of a deposit as part of the circulating medium. Most persons would accede to the statement that a check serves to effect payments as well as does a coin or a paper note; but they would say that it is the check, not the deposit, which serves as money. Yet a moment's reflection shows that the check bears the same relation to the deposit as the coin actually used in making payments bears to the coin carried in our pockets. Both are for potential use. Not all the coin (taking coin as typical of money that passes by mere delivery) is buying commodities all the time. Part of it is being carried in pockets or kept in tills, as a fund to be used and paid out at convenience. The portion of it used at any time in purchases is only part of the total then existing and available. Deposits similarly make up a fund, an available means of payment, to be drawn on at convenience. When counting up the total quantity of coin in a community, we count the whole supply on hand, not merely that which happens to be making purchases at a given moment. Similarly in reckoning this other form of circulating medium, we must count up the total volume of deposits, not just that part which happens to be in use at the moment in the form of checks. The check is simply the deposit in actual use; it is the deposit in motion; it is the deposit which is passing from one person to another.

Bank deposits are often spoken of as "credits." They may equally well be called debts. They represent claims which the depositors hold against their banks—rights to collect legal tender on demand—and from this standpoint they are credits. From the opposite standpoint they are obligations of the banks to their depositors, mere debts, promises to pay out legal tender on demand. The movement of deposits from person to person, then, is a transfer of what may be regarded either as the credits of the depositors with the banks or as the debts of the banks to the depositors. In this respect bank deposits differ from other kinds of money, and more particularly from legal tender money. "Legal tender" means that the money must *ipso facto* be accepted in discharge of a debt. Bank deposits cannot so operate. Before a person will accept payment by check he must first have some assurance how the drawer stands with the bank upon which the check is drawn, and assurance also of the ability of the bank to honor the check when presented. And if he is well informed about the law, he may bethink himself that the bank is under no obligation to him,—to the payee. Its only obligation is to its depositor, the drawer; and even here the responsibility is only for damage accruing from refusal to honor the check (e.g., because of injury to the depositor's standing or general credit in the community). Payees must be on the watch, and so must banks.

These circumstances—the intermingling of credits and debts, the somewhat intricate legal position of banks and depositors and payees—might be expected to stand in the way of the wide use of deposits and checks, and their easy and smooth operation as money. Obstacles of this kind do still impede their use in many countries, as France, Italy, even Germany. But where people have become habituated to deposit banking, the difficulties have become unimportant. Banks do indeed keep a watchful eye for "bad" checks (where the depositor has no balance) and counterfeited checks; a check proffered by a stranger is subject to scrutiny and calls for explanation. Practically all checks, however, come from persons who are known or vouched for and they are accepted unhesitatingly as "cash." Fraud and the possibility of non-payment are ignored as much as is the danger of violence and robbery

on the streets; our daily doings are quite unaffected. The system, elaborate and complex tho it is, works with extraordinary smoothness and despatch.

Deposits in their working as a circulating medium are among the most marvellous of economic phenomena. Like the division of labor which they serve to facilitate, they have developed by no intention, and until very recent times with but little restraint or guidance from legislation. They work out their results by processes only half understood by the very persons who manage them, the bankers. They combine in a remarkable degree safety and convenience. They are safe from theft because a check is payable to a specified person, and the bank is answerable for making payment to that person only or his endorsee. They are convenient because a few strokes of the pen serve to remit any sum however large, and remit that sum precisely to the last cent. Hence they have become, in the countries where the system has reached its full development, the dominant factor in the circulating medium.

Because of the peculiar character of deposits, some writers do not include them in "money" at all. Whether to do so is a matter of choice in terminology. The fundamental thing is that all the various means of payment circulate and perform the functions of a medium of exchange; and in this essential respect bank deposits act precisely as do coin and notes. As regards the fundamental problems—the value of money, the price level, the stability of prices—they are not only important but of cardinal importance.[1]

§ 5. What now determines the value of money? That is, what determines the general range of prices?

The first step toward answering this question is to understand the special relation between the quantity of money and its value. The relation is fundamentally a very simple one. Double the quantity of money and, other things being equal, prices will be twice as high as before and the value of money one half. Halve the quantity of money and, other things being equal, prices will be one half what they were before and the value of money double.

[1] It may be noted that only those deposits habitually used for making payments —typically the deposits subject to check—are to be included. Savings deposits not subject to check are of a different character.

That an increase in quantity tends to lower value is a proposition holding good of all commodities. The special proposition concerning money is that its value tends to vary in precisely inverse proportion to its quantity. This exact kind of relation does not hold good of any other commodity. Double the quantity of wheat, and its value will probably fall to much less than half of what it was before. Double the quantity of sugar, and its value will probably fall by no means to one half. For both wheat and sugar the outcome will depend on the elasticity of demand. But in the case of money there is no doubt about what the elasticity of demand is, and no such difficulty in prediction. The value of money, under the simplest conditions, is exactly inverse to its quantity.

Such is the so-called "quantity theory" of money. Concerning it a hot controversy has long been waged. It has been vehemently denied, and often it has been erroneously stated. Rightly stated it is not to be questioned, but it must be understood as well as rightly stated. In the preceding paragraph it has been put boldly, with the purpose of bringing out clearly the fundamental truth. But the reader will note the phrases "other things being equal" and "under the simplest conditions." Great qualification and elaboration will be required before the bald statement can be made to fit the complicated phenomena of modern life. The last word cannot be said until a long series of topics have been covered. For the present, let us consider the essential ground on which the proposition rests, and some of the important qualifications.

These essential grounds are found in the nature of the demand for money.[1] People often say that the demand for money is without limit. They mean thereby that any individual desires to secure possession or control of all he can. But he desires possession or control as a means, not as an end. Money is not directly en-

[1] "Demand" is used here in a different sense from that in which the term was used in Chapter 10, § 1. The demand for money, as spoken of here, means the quantity of commodities of all sorts which, being put on sale, are offered for money. Ordinarily, when speaking of a particular commodity and of the demand for it, economists mean by "demand" the quantity of that commodity which is *demanded*, not the quantity of another thing (money) which is offered for it. It is in this sense, of quantity demanded, that we construct the "demand curve" for a commodity. But as regards money we speak of demand in the other sense, more familiar in everyday usage: what is offered in exchange for it.

joyed, not eaten nor drunk. It is sought in order to be spent,—as a means of getting other commodities. We may set aside as negligible the case of the miser who gloats over money for its own sake, and also some other possible cases of hoarding. All the money, whether any individual has control of much or little of it, is spent sooner or later. The demand for it—what is offered in exchange for it—consists of the commodities on sale. But the commodities on sale are simply all the commodities that are to be exchanged. The demand for money, in any given community at any given time, is *constant*,—constant in the sense that it is not subject to variation because of the greater or less range of prices. Whether goods sell for less or more, all of them will still be sold, and will still be offered for money. Hence when there is twice as much money, the same number of commodities will be offered for the money, and prices will be twice as high as before.

In other words, using a phrase already explained, the elasticity of demand for money is unity. Herein the position of money is unique. As regards the immense majority of commodities, demand is elastic in some cases, inelastic in others, but rarely so balanced that the same sum is always spent on any one. The case of money is peculiar in that the total amount of goods offered in the market is not affected by its value. The total remains always the whole number of commodities that are exchanged. The total may indeed change; more of commodities may be produced, and more may then be offered in exchange for money; but more are not so produced and offered *because* the value of money is less. Extraneous causes, in this case as in others, may lead to a new situation,—a change in the price level. But given the same population, the same output of goods by that population, the same ways of selling and marketing—and this is what is meant when we say "other things remaining the same"—the demand for money is a constant amount.

This characteristic of the demand is the result of the very nature and uses of money. The elasticity of demand for cotton or for apples may happen to be unity. Some inquiries on the fluctuations in the prices of cotton have indicated that in fact such is approximately the situation, at least within certain ranges of supply

and price. But no one could predict it in advance; whereas a consideration of the very nature of money and of the uses which money serves, points to the conclusion that the demand for it is necessarily of this special character. The conclusion would not hold good of the precious metals when used for other purposes than coinage. If the demand for silver plate or platinum jewelry should prove at a given time to follow the same course, we should be interested, but surprised; there is no *a priori* reason for expecting the phenomenon. But in the case of money we cannot be surprised; the result is what must be expected.

§ 6. Let us now begin to introduce the explanations and qualifications of this fundamental principle. In the first place, we do not have in mind the whole number of commodities, nor even the whole number exchanged, but only the number exchanged thru the medium of money. It is these last alone that constitute the demand for money. Some goods are consumed by those who produce them and do not enter the circle of exchange at all. Such are agricultural products consumed by those who grow them. These evidently do not constitute at any time demand for money. With the growing elaboration of the division of labor, the proportion of goods so used tends to become steadily less. In a country like the United States at the present time it is not far from the truth to say that all things that are produced are exchanged.

Nor is it far from the truth to say that the exchange of things takes place solely thru the medium of money—to say that all things exchanged are sold for money and are thus exchanged thru money. True, there may be barter. The farmer may bring his eggs or grain to the country store, receive credit on the books of the dealer, and subsequently "buy" goods in terms of money, the amount being debited to him. Here the transaction, tho it is in terms of money, essentially is one of barter. The volume of transactions of this sort may be not inconsiderable in the United States, yet it is small in comparison to the total of transactions. Barter, such as this is, has disappeared even more than production for one's own consumption.

§ 7. More important for the main line of reasoning is a qualification as to the rate or manner in which goods and money meet

each other in exchange. The preceding statements seem to imply that all the goods are exchanged for all the money in one transaction. Obviously this does not happen. At any given moment, or on any given day, only a fraction of the goods is being sold and only a fraction of the money is being used in purchases. Here, as elsewhere in economics, we should have in mind a flow, rather than a fund. The total stock of commodities is indeed sold sooner or later. But only a portion of it actually comes to market in any one day or week or other unit of time, the rest following in more or less orderly sequence. There is a flow of goods into actual exchange. Similarly, the total quantity of money may be regarded as a fund, while as actually used for purchasing goods it enters the market as a flow.

The phrase "rapidity of circulation" or "velocity of circulation" of money has been used to indicate this obvious fact. Of the total money actually on hand in a community a portion only is at any given time at work, so to speak. The money idle in our pockets does not directly influence and affect prices; only that which is buying goods at the counter does so. What proportion is on the average at work depends on the habits of the people. It is affected by their geographical distribution and by the character of their industries. In a thinly settled agricultural section, where access to shops is not easy or frequent, a larger portion of the money is likely to be idle than in a thickly settled manufacturing or commercial section. And the temper of the people is a factor. If they are confident of themselves—perhaps unduly confident, and thoughtless of the morrow—they are likely to spend money as fast as it comes into their hands, and little of it remains idle at any time.

These remarks apply to the larger transactions of merchants and dealers as well as to the everyday purchases of consumers. Traders and producers always have on hand more money than they are using in purchases; the proportion depending partly on the nature of their business operations, partly on their temperament. The fact that in countries like the United States these classes use not actual cash but checks against bank deposits does not alter the situation; it only supplies another illustration of the difference be-

tween the fund of money and its flow. The total of their deposits in banks constitutes the fund; the checks by which purchases are effected from day to day constitute the flow. In every form, the medium of exchange has its flow, or rate of use—its rapidity of circulation.

Similarly, goods have their rapidity of circulation. In more familiar language, they have their rate of turnover. This also depends on a great variety of circumstances. It is likely to be rapid in a large city, where merchants can gauge fairly accurately the power of their markets to absorb goods; it is likely to be slower in the country where, owing to the smallness of the number of buyers, demand is hard to estimate in advance. It is affected, like the flow of money, by the temper of the people. In an energetic and restless country like the United States the turnover is likely to be quicker than in a comparatively staid country like France. It varies, too, in different branches of trade. The turnover of a grocer's shop is more rapid than that of a hardware dealer's, that of a flour mill quicker than that of a textile factory. In ordinary times the flow of goods as a whole takes place steadily and continuously, and in a given community with a surprisingly regular course.[1]

The proportion of money which is actually buying goods is not accidental; it is determined by the silent force of custom. It may be irregular for an individual, but over thousands and millions of individuals it follows a steady course. The flow of goods to market takes place in general at a similarly regular rate.

Hence we may expect with some confidence that if the total quantity of money be increased, that quantity which is used in making purchases at any given time will be correspondingly increased. Suppose, for example—to use an illustration of Mill's—that suddenly everyone in the community has twice as much

[1] In times of disturbance and in the several phases of business cycles, this statement would have to be qualified; how qualified and how much, it is not easy to say. We know less about rapidity of movement for goods than for money. With goods it is a matter of inventory and turnover: the relation between the goods in stock at a given date and the period during which these are sold and have to be replaced by a new stock. This relation is probably more constant for retail trade than for wholesale trade, and more constant for wholesale trade than for manufacturing. In periods of depression, inventories are likely to be lower all around than in periods of activity. Over long periods, easy access to sources of supply thru the telephone, railway, motor probably tends gradually to make the trend of inventory holdings decline.

money. The only thing that can be done with it is to spend it. Once the change is established, there is nothing to alter the habits of the people; nothing to cause a larger proportion to be kept in the pocket or in reserve. The quantity of goods remains the same, nor is there anything to alter the mode in which people and dealers bring their goods to market. The flow of money will be doubled, the flow of goods unchanged, and prices will be twice as high as before.

The same effect which would ensue from a doubling in the quantity of money would ensue also from a doubling of its rapidity of circulation. If twice as large a proportion of the total stock is steadily in use for purchasing goods, the effect is the same as if the quantity were doubled without any change in the ways of using it.

The propositions which were laid down in the opening section of this chapter obviously assumed that the quantity of goods and the flow of goods into exchange remain constant. This is what was implied by the qualification "other things remaining the same." Needless to say, the quantity of goods does not always remain the same. If it be doubled when the quantity of money is doubled, prices will be unchanged. If goods be doubled, money remaining as it was, and the flow of goods to market unaffected, prices will fall one half. If the flow of goods to market—their rapidity of circulation—be so affected that twice as large a proportion of goods is steadily offered, prices will again fall one half.

Rapidity of circulation is greater for money than for goods. To put it in other words, the proportion which at any one time the money actually offered for goods bears to the total supply of money is greater than the proportion which the goods offered for money at any one time bears to the total supply of goods awaiting exchange. The reason for this difference is obvious. Money can always be used without delay in purchases; goods can be sold but slowly. Money need never wait for a buyer; goods must often wait for one. Many commodities have necessarily a slow turnover, as hardware and household furniture. Money comes into the market quickly. Tho there may be hoards, and occasionally an accumulation of unused money in the hands of people who are getting

larger incomes than they are used to, money in the main is kept at work briskly, at a rate which for any given time and country is determined by the ways and customs of the people.

§ 8. Our first proposition then—the main thing for the immediate purpose—is that the level of prices tends to vary in proportion to the quantity of money.

If this were the whole story, changes in the volume of money would be of no consequence. Greater or less quantity would alter the number of counters used in effecting exchanges of goods, but it would do simply this and nothing more. The same amount of goods would be produced as before, the same number of exchanges would be carried out, the same level of real income and the same distribution of real income would prevail. The change in money would do nothing to affect the economic well-being of the community. Problems concerning the causes governing the value of money might remain of intellectual interest, but they would be of no practical importance.

So much is true, *if* this were the whole story. It is by no means the whole story. While it is of no consequence whether the general price level, when settled, is at one height or another, a process of change upward or downward may have substantial effects of the first consequence. That they are important has long been perceived; and it has been borne in on us with renewed vividness by the extraordinary disturbances of the twentieth century. And these disturbances have brought also a better understanding of the whole set of problems raised by changes in prices. The theory of money and prices needs to be elaborated into something very much less simple than that which has been presented in this chapter; so different that not a few modern critics would discard the simple introductory proposition once for all, as giving a quite inaccurate picture of money and its working in the world as we have come to know it. The debates on the "quantity theory" have been acrimonious, even tho they have turned largely on matters of exposition and emphasis. They do indicate the need of enlargement and correction of the first general statement. It will be the object of the chapters that follow to meet this need.

§ 9. What, now, is the total quantity of money? Since this con-

cept will be used repeatedly in the subsequent chapters, its meaning must be noted carefully. In discussion of the bearing of the money supply on the price level, we mean by quantity of money the total volume of all forms of means of payment—coin plus notes plus deposits—in the possession of the public. It is the quantity of money *in circulation.* It includes all coin and notes in people's pockets, in their purses and cash boxes, in the tills of business establishments. It includes all bank deposits against which checks may be drawn. It does not include deposits in savings banks, "savings deposits" and "time deposits" in commercial banks against which checks may not be drawn, or deposits by one bank in another. It does not include, further, coin or notes in the possession of banks which are held for meeting withdrawals by check. Banks do not hold such cash for use in any current purchases. They hold it for meeting withdrawals by their customers, and it becomes part of the volume of active money only when withdrawn by the depositors. To count both the deposits of the customers and the cash held as reserve against them by the banks would be to count the same thing twice. This applies also to the huge amounts of specie or coin which Central Banks and Government Treasuries hold as backing or reserve against deposits and notes. There are some nice distinctions to be drawn with regard to such holdings, both those of ordinary banks and those of public and semi-public banks; but they are not important for the present purpose. We fix attention here on the active volume of money, the sum total of the net means of payment *now* available for immediate use. It is this with which we have to deal in the chapters that follow.

CHAPTER 20

CHANGES IN PRICES

§ 1. Changes in the price level are measured by index numbers. The arithmetic mean. Illustrations from United States prices, 1913–18.—§ 2. Other means: the geometric, the median. Weighted index numbers.—§ 3. Prices and incomes, money; debtors and creditors.—§ 4. Complications. Changes in particular prices ordinarily are of chief direct concern. As regards the price level abrupt changes alone have far-reaching consequences.—§ 5. Debts having a long time to run. The problem where the general trend of prices is different from that of money incomes.—§ 6. Prices and costs. Effects of changes in their relation resulting from the nature of capitalistic industry. Business concerns and business men as debtors and creditors.—§ 7. Prices and costs, *continued*. How far their relation is unevenly influenced by industrial oscillations.—§ 8. Wages as costs and their relations to prices. Similar problems with monopolized raw materials.—§ 9. Concluding remarks. The commodity dollar.

§ 1. Two topics will be taken up in the present chapter: first, how to ascertain and measure whether changes in prices have taken place; second, some of the salient consequences for good or ill of such changes. Of the causes of the changes nothing more will be said for the present.

The measurement of changes in the value of money would be easy if all prices went up and down together. But this they never do. Some prices go up, while others go down. Occasionally, in times of very great and rapid movement, almost all prices change in the same direction. Even then, some rise or fall in less degree than others. In 1916–18 the general trend of prices in the United States was unmistakably upward; most commodities advanced sharply in price, some enormously. Yet a few showed a downward movement. The extreme range was from oil of lemon, whose price in 1918 was only one third of what it had been a few years before, to a drug, acetophenetidin, whose price became fifty times as high as before.[1] Tho the fact of a change in a given direction be clear,

[1] I take these illustrations, and also those on page 262, from the *History of Prices during the War*, prepared by Professor W. C. Mitchell and published in 1919 by the War Industries Board, Washington.

the complexity of the phenomena makes difficult the measurement of its extent.

To get a summary expression of the general trend of prices, resort is had to the method of index numbers. An example will best explain how an index number is constructed. Suppose that on January 1, 1900, the price of iron was $15 a ton, of wheat $1 a bushel, of cotton 10 cents a pound, of wool 40 cents a pound. These are called the base prices. Later prices are expressed in relation to them, usually by stating them in terms of a percentage. Suppose that a year later, on January 1, 1901, the prices of these four commodities have come to be $20 for iron, $1.25 for wheat, 10 cents for cotton, 36 cents for wool. Then the actual prices, and the percentage relation between them, would stand thus:—

	1900		1901	
	Base Price	100	Price	Percentage to Base
Iron	$15.00	100	$20.00	133
Wheat	1.00	100	1.25	125
Cotton	.10	100	.10	100
Wool	.40	100	.36	90
Total		400		448
Average (arithmetic mean)		100		112

The index number was 400 for 1900, and rose to 448 for 1901. Reduced to the arithmetic mean, the index number for 1900 was 100; that for 1901 became 112. The base average, of course, is always 100; the average for any other year is then a percentage of the base average. In the example just given, the index number shows a rise in prices of twelve per cent; or, rather, as the very word "index" implies, *indicates* a rise to that extent.

If, now, instead of four commodities, fifty or a hundred were treated in this way, we should feel some confidence in the indication obtained as to a general change in prices. If the summarized result for a large number of articles is an advance of ten or twenty per cent in the index number, it is tolerably certain that most

commodities have gone up in price. No doubt the result may be ascribable to the fact that half the commodities went up a great deal, while the other half went down, tho but moderately. But an examination of actual changes, even a cursory one, almost always shows, where a marked change has occurred in an index number, that the large majority of prices have moved in the one way indicated. The index number serves, therefore, to point to a fact—that on the whole prices have moved in one direction.

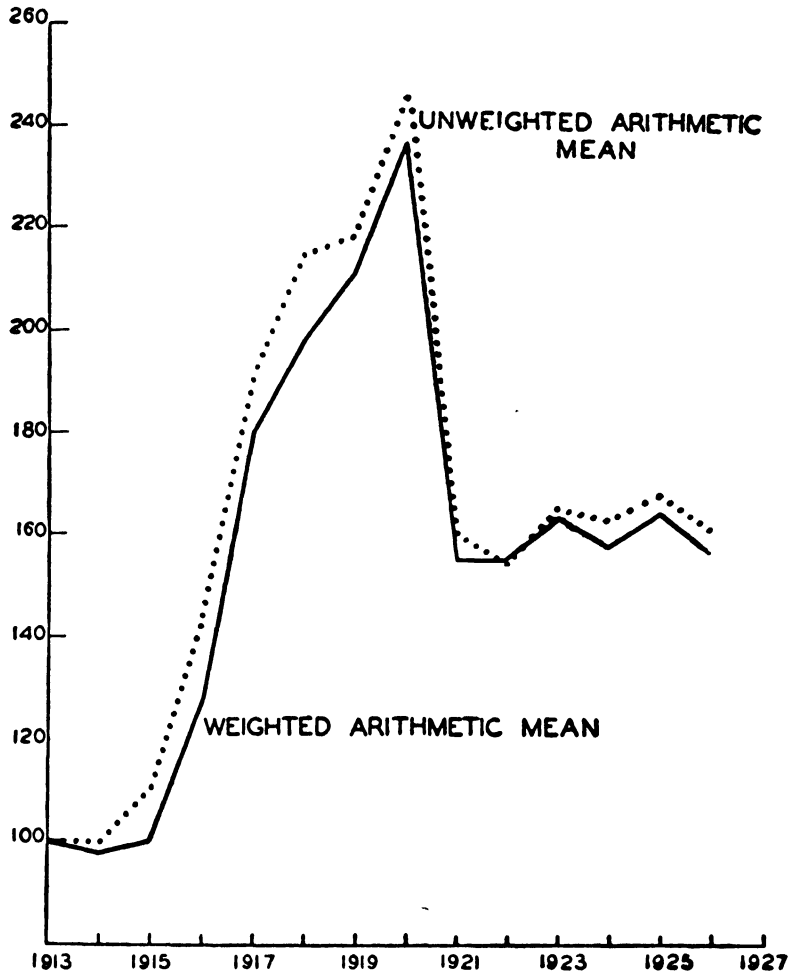

FIG. 16

To illustrate: during the years immediately preceding the war of 1914–18 and during the war period itself, the index numbers for prices at wholesale in the United States, calculated on this plan, were as given at the top of the next page.

July 1, 1913 to June 30, 1914 (base)	100
Calendar year 1913	101
" " 1914	99
" " 1915	102
" " 1916	126
" " 1917	175
" " 1918	194

A sharp advance in prices is here indicated. It is true, as has been mentioned, that a few commodities actually fell in price; yet they were no more than two per cent of the whole number of articles. With these almost negligible exceptions, all commodities sold at higher prices in 1918 than in 1913–15. Yet, unmistakable as was the general trend, nothing like a uniform movement took place. Grouping the several articles according to the extent of the advances, we find that in 1918

9	per cent of the articles showed prices between 130 and 149
15	" " " " " " " " 150 and 169
15	" " " " " " " " 170 and 189
11.5	" " " " " " " " 190 and 209
9.5	" " " " " " " " 210 and 229
7.0	" " " " " " " " 230 and 249
67.0	" " " " " " " " 130 and 249
10.3	" " " " " " " " of less than 130
22.6	" " " " " " " " of 250 or more

In other words, two-thirds of the articles were sold at prices ranging from thirty per cent higher to one hundred fifty per cent higher. The change, tho it showed great variety and irregularity, was almost universal; the index number, which registered an advance of ninety-four per cent, or almost a doubling of prices, summarized an upward movement which was great and rapid, and yet was highly complex.

§ 2. Other modes of reaching index numbers have been proposed, the arithmetic mean being criticized as crude and inadequate. Some of the suggested improvements may be briefly noted, and the usefulness of the simpler method tested by comparison with the results from those more complex.

The geometric mean has been advocated; and sometimes other mathematical means. Of the geometric mean it is said, with undoubted truth, that its use will mitigate a misleading effect on

the index number from extraordinary fluctuations in the price of a single article. With the use of logarithms the geometric mean is easy to ascertain; and it has quite as good a right to be entitled a "true" average as the arithmetic.

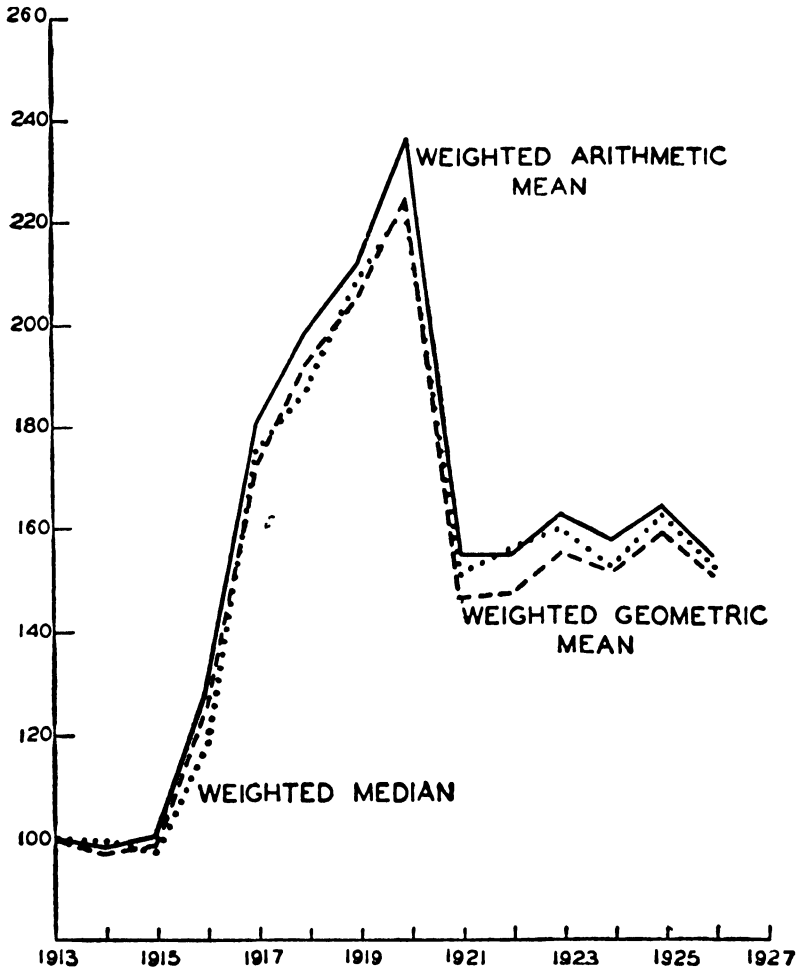

FIG. 17

Another proposal is for the use of the median. Let the index numbers be made up, not by averaging, but by ascertaining midway points. Arrange the several price quotations for any year (reduced to a uniform basis as for the other methods) in numerical order, and then ascertain that figure which stands in the middle of the series—that figure on either side of which there is an equal number of quotations. For various sorts of observations the median is thought by statisticians to be at least as significant as any aver-

age; and tho comparatively unfamiliar, it is easy to use. Even more than the geometric mean, it prevents an extremely high or low price of some one article, or of a very few articles, from having an undue influence on the index number.[1]

Different is the improvement of the simpler method itself—the arithmetic mean—by taking account of the relative importance of the different articles; or, as it is technically put, by *weighting* the articles. A change in the price of wheat, for example, is of much more importance than a change in the price of wool. If wheat were to double in price, the purchasing power of a given income would be seriously affected; if wool were to double in price, much less. The varying importance of different commodities may be regarded in the construction of an index number by assigning weight to the commodities in the proportion of their consumption. If the community as a whole spends four times as much of its income on wheat as on wool, wheat may be counted as if it were four articles and wool as if it were one. If twice as much is spent on cotton as on wool, cotton may be counted as if it were two articles; while iron, on similar assumptions, may be counted as three. The prices used in the original illustration would then be made up into an index number as shown in table on page 265. This weighted average indicates a rise in prices from 100 to 119, whereas the simple average would have indicated one from 100 to 112 only. And the weighted average is plainly the more significant; since the higher prices of widely used articles like wheat and iron are more important than the lower price of the less-used wool.

Tho the weighted index number is clearly preferable, the application of this more refined method presents difficulties. It is not

[1] Thus if a series of price quotations, reduced to a basis of 100, were

86	102
90	106
94	110
97	150
100	

the median would be 100. If the last figure were not 150, but 120, the median would still be 100.

There being in this series an odd number of figures, the median is *the* middle one. If there were an even number, the median would lie between the two middle figures, and would be in so far indefinite. But where there are many figures, as is always the case with price quotations, the median is sufficiently precise.

easy to ascertain the consumption or relative weight of the several articles, especially where a very large number (100 or more perhaps) are included in the list. Moreover, the consumption of the different articles varies. Changes in habits take place; a given article may be much less used in 1910 than in 1900; how readjust the weight given it and the whole weighted index number? These difficulties, and others that might be instanced, tho not insuperable, add to the complications of weighting.

In regard to all these suggestions, whether for improvement in the arithmetic mean or for the use of a different mean, it must be borne in mind that no index number corresponds to a real thing. It is not like the mean of certain observations in natural science

	1900			1901		
	Weight	Base Price	Weighted Base	Price	Percentage of Change in Price	Weighted Change in Price
Wheat	4	$1.00	400	$1.25	125	500
Cotton	2	.10	200	.10	100	200
Wool	1	.40	100	.36	90	90
Iron	3	15.00	300	20.00	133½	400
Total	10		1000			1190
Average			100			119

—such, for example, as those for measuring the distance between the earth and the sun—of which any one may err, but whose average will point to a specific fact. An index number points to no specific fact. It gives, to repeat, only an indication of the general trend of prices. People often speak and think loosely on this topic, as if an index number told the whole story once for all. There is no one change in prices but a medley of many, different in direction and degree. All we can hope to secure by averaging and summarizing is some concise statement of the general drift.

Now experience in the application of the various methods to the same sets of figures shows that the simple arithmetic mean, when applied to a sufficiently large number of price quotations, gives substantially the same results as more refined methods. If many articles are in the list, some of much importance, some of

little, it is unlikely that all the important articles will fluctuate in one direction and all the unimportant in another. If they did so (as in the example just given), weighting would be indispensable. But the fluctuations in fact are likely to be distributed among the several classes in much the same way. An unusual change in the price of a particular article, whether it be consumed in large amounts or in small, will not affect greatly an average made up

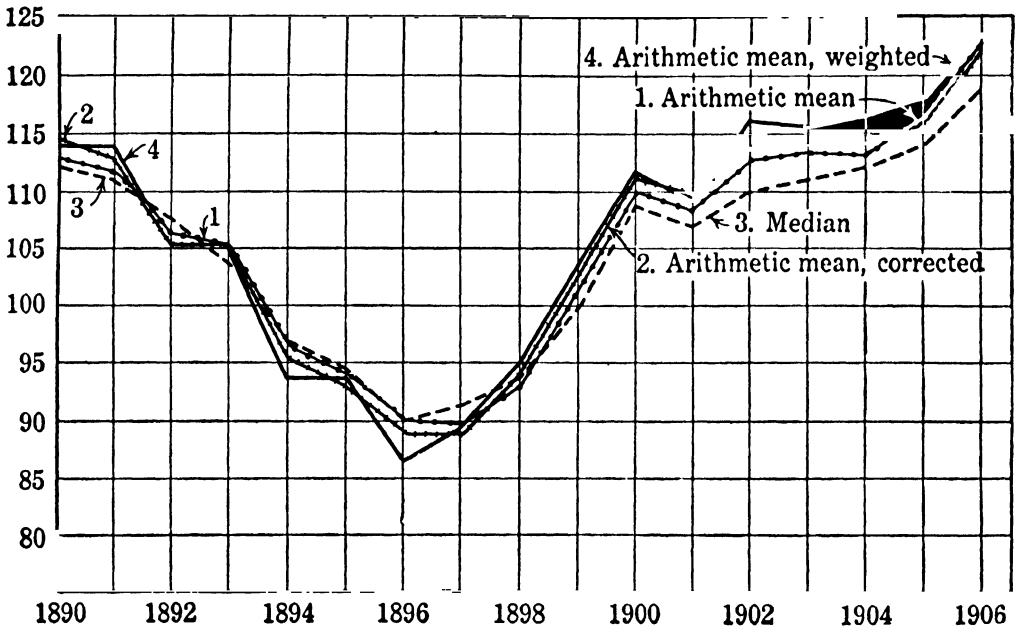

FIG. 18

from many price quotations. And in practice it has been found that the simple unweighted average brings results not very different from those obtained after weighting. Similarly it has been found that the method of the median does not yield results substantially different from those of either the simpler or the weighted arithmetic mean.

This similarity of outcome is illustrated by Figure 18, showing the course of four index numbers reached in different ways, all based on the same quotations of prices.[1] One represents

[1] The four series are:—

(I) The Department of Labor's arithmetic means, for prices of 250 articles.

(II) Professor W. C. Mitchell's rearrangement of the same price figures: "The Bureau's list of commodities contains anomalies such as the inclusion of a single series [of quotations] for wheat and ten for cotton sheetings; two for hogs and three for glassware, etc. The result is most unscientific weighting in what purports to be an unweighted index number. To remedy this obvious defect, I have combined the

the simple arithmetic mean of 250 price quotations; the second, another arithmetic mean of the same prices consolidated into 145 quotations; the third, the median of these same 145 quotations; the fourth, a weighted index number of 50 among these commodities. The prices are at wholesale, in the United States, for the period 1890–1906; the "base," indicated by 100, is in each case the average (arithmetic mean) for the decade 1890-99.

§ 3. Nothing has been said in the preceding section about money incomes in their relation to the prices of goods. Money incomes are obviously important elements in the price situation; yet they are not considered in making up indices of change such as have just been considered. Yet the relations between money incomes at different times are at least as important as those between prices of goods.

series for nearly identical articles, thereby reducing the number of series to 145."—*Journal of Political Economy*, May, 1910, p. 372; compare the same writer's *Gold, Prices, and Wages under the Greenback Standard*, p. 19.

(III) The median for the same (145) series of quotations, as calculated by Professor Mitchell.

(IV) A weighted index number for 50 staple articles, selected from among the 250 (145); the weighting being on the plan of the Gibson index, but revised by Professor Mitchell.

The figures of the four series are:—

	I Arithmetic Mean of 250 Quotations	II Arithmetic Mean of 145 Quotations	III Median of 145 Quotations	IV Weighted Index Numbers from 50 Quotations
1890	112.9	114.1	112	114.0
1891	111.7	112.7	111	113.9
1892	106.1	106.1	107	105.1
1893	105.6	105.0	104	105.2
1894	96.1	95.6	96	93.9
1895	93.6	92.8	94	93.9
1896	90.4	88.8	90	86.6
1897	89.7	88.7	91	89.2
1898	93.4	93.5	94	95.0
1899	101.7	102.5	100	103.4
1900	110.5	111.3	109	111.6
1901	108.5	109.6	107	109.2
1902	112.9	113.7	110	116.2
1903	113.6	113.8	111	115.3
1904	113.0	113.9	112	116.3
1905	115.9	115.8	114	117.9
1906	122.5	122.3	119	123.4

The most familiar case in which incomes must be considered is that of debts and contracts,—engagements for the future. Let us first consider a case quite simple, a mere change in the value of money: all things go up or down in the same direction and to the same extent; not only prices of goods but also money incomes of all kinds. So far as concerns operations which are strictly contemporary, no one is helped or harmed. The only change is that more units of money are used in every payment. But the essence of debt —of borrowing—is that there are two operations separated in time: first borrowing, then repaying. The borrower receives from the lender at one time, the lender is repaid by the borrower at another. Both do so in terms of money. A fall in the value of money —a rise both in prices and in money incomes—is then of advantage to debtors; the converse change is of advantage to creditors. When prices and incomes go up in the interval between the contracting and the paying of a loan, the debtor, on returning to his creditor the amount of money borrowed, returns less in the way of commodities. This is the gist of what happens; tho what the debtor is chiefly conscious of is that his debt is disposed of with comparative ease, his money income having become larger. Conversely, when prices and incomes go down, the debtor, on returning the same money, returns more in the way of commodities; and, having less income, he finds it comparatively hard to pay the debt. In effect the changes in the value of money cause a shift in the distribution of wealth, and one which on its face is related to no principle of equity or of social justice.

§ 4. So simple a change never occurs. As we have seen, all prices do not go up and down together; neither do all incomes. Each individual is chiefly concerned with the particular articles bought or sold by him, and with their prices. For short periods—two or three years—the price of any one article may show no clear relation to the general price level. It may go up and down, regardless of the general drift indicated by an index number. That general drift, again, often is not easy to make out until it has been going on for some time. Under usual and "normal" conditions, changes in the general price level take place slowly, and in the transactions of everyday life are hardly noticed. Even from year to year, there

is ordinarily little change that attracts attention; the observations of everyday life must extend over several years if we are to be impressed by the fact that any general rise or fall is really in progress. Most debts, however, are for moderate periods. A change in the price level of 5 per cent a year or even 10 per cent, as registered in an index number, would be little noticed by debtors or creditors. What concerns them forcibly and immediately is a change in the prices of the commodity or commodities in which they deal, and these prices may not be closely related to the general movement and often are more abrupt and noticeable than that movement. All things considered— the slowness of changes in the price level and the immediate importance of particular price movements—the disturbance of the debtor-creditor relation by the movement of price level is not for most debts of so much moment as would appear at first blush.

Quite different of course is the situation where there are great abrupt changes in the price level, such as take place when paper money is put out in huge volume. Then the general movement upward is so vast that the peculiarities of individual price changes become of negligible consequence. Creditors may be virtually compelled to turn over their wealth to debtors. The injustice becomes unmistakable. On this dreary subject more will be said in a later chapter.

Something similar happens, tho rarely with the same dramatic suddenness, when an abrupt fall in prices takes place. Debtors then are clearly seen to suffer at the expense of creditors; if not so overwhelmingly as in cases of violent inflation, yet with profound distress. Such distress, however, has never resulted directly from legislation; no country has passed laws which directly caused catastrophic results of this kind. A great and rapid fall in prices does not come except from some unexpected economic collapse; possibly when a war breaks out, more often as the aftermath of an economic crisis. In the last-named case, the fall is more or less acute according as the overturn from excitement to depression is quick and severe. When this happens, the consequences are less uniform and less widespread than in the case of inflation. Some debtors feel them more than others, according as the goods they

produce are more or less affected. After the crisis of 1929, for example, the prices of agricultural products fell abruptly, not merely because of the general collapse but because they were pushed down by the peculiar conditions of international trade that ensued in almost all countries.

§ 5. The case is somewhat different when we consider not short-time or abrupt changes but debts having a long time to run. As regards these, even under the comparatively stable monetary conditions which we think of as usual or normal, there may be much injustice and hardship. The longer the time, the more is this likely. In the course of twenty years, possibly in the course of ten, marked changes in the range of prices may occur, and with them marked injustice to debtors or to creditors as the case may be. The longer the period, the greater the likelihood of harm. Tho obligations running very long periods are not often contracted by individuals, they are common on the part of corporations and of governments. European governments, to be sure, when they borrow, usually do not undertake to repay the principal sum at any given date. They promise only the regular payment of a stipulated rate of interest; and while they reserve the option of repaying the principal either at times expressly stated or at their discretion, they do not undertake to repay unless it suits them. In such case they have a protection against loss from price changes, but their creditors have none. In the United States, the government has often borrowed on long time and exposed itself to possible loss; a practice, however, which has been kept within limits moderate enough not to cause great difficulties. Many of our great corporations, however, and especially the railways and the public utilities, have borrowed quite without regard to possible price changes, and not infrequently with complete disregard of possible changes in the rate of interest. Bonds have been issued payable after the lapse of forty, fifty, and even one hundred years, without provision for redemption or interest adjustment in the interval. Who can say what will be the range of prices after the lapse of a century?

Such long-time obligations find a market because most investors (like other people) think of the value of money as unchanging,

and because they are glad to have a long-time guarantee of an income supposed to be fixed. And it is inevitable that borrowers, when they wish to raise great sums of money for permanent investment, should adopt the devices which will attract the investor. In such engagements both debtors and creditors take great and unpredictable risks. Under monetary systems as they now are, and are likely long to remain, these risks can be reduced to the minimum only by restricting all loans to periods of a moderate number of years.

§ 6. A different question of justice between debtor and creditor arises from the fact that money wages and other money incomes do not necessarily move in the same way as the prices of commodities. In the preceding sections, it has been tacitly assumed that these two movements —of prices and of money incomes—are parallel. But one may lag behind the other; or the movements may be in opposite directions. Here we deal with long-period changes. They are of quite a different sort from those considered in the preceding section, and raise questions of another order.

Suppose, for example—to take the sort of case which fortunately is most probable—that industry is progressing, the arts are advancing, the prosperity of the community growing. This means that real incomes are becoming larger; that the commodities at the command of the community as a whole, and on the average for each person, are more abundant. The concrete way in which that abundance must show itself, where all transactions and all exchanges are carried on through money, is in cheapness of goods relatively to incomes. Goods may become cheaper, money incomes remaining the same; or money incomes may become greater, prices remaining the same; or some intermediate relation may appear. In any case, prices and incomes will not move together. Relatively to prices, money incomes will rise.

Thus, during the period of falling prices after 1873, money incomes on the whole did not fall. The evidence to prove this relates chiefly to the familiar crafts and to unskilled or little skilled labor; since comparison of wages at different times is here easiest. Money wages on the whole did not fall after 1873; they rather tended to rise. So it was with the rates of wages which are eu-

phemistically called salaries—the pay of teachers, corporation employees, public officials. The same upward tendency or, at the least, stationary tendency showed itself also in the more irregular money incomes of professional and business men. With rising or stationary wages and incomes all around, and with falling prices, real incomes in terms of commodities and of services must have gone up substantially. Obviously this is the natural outcome of industrial progress and cheapened production. That same outcome of progress and cheapness must be expected to appear in a period of rising prices; only in this case in a different way. If prices advance, money incomes must advance at least as much, if real income is to remain the same. And if fundamental forces are at work to promote progress in industry, wages and all money incomes must advance even more than prices. If in the future an increasing supply of money acts to bring about continuously rising prices, we must expect that this change will be accompanied in time by an even greater rise in money incomes.[1]

What, under such circumstances, are the relations between debtors and creditors? With prices falling and incomes stationary, debtors, paying their debts with the same amount of money, repay to creditors more in the way of commodities. This may be called repayment according to a labor standard. It is true that the debtor pays back more commodities than he got; but those commodities represent the same money income and (presumably) the same amount of labor as before. It may be argued fairly that the debtor suffers no injustice if at the time of repayment he has the same money income as when he contracted the debt. The creditor simply shares in the greater cheapness of commodities due to improved production. Suppose, on the other hand, that there are stationary prices and rising incomes. The debtor, paying back the same money, pays back also the same commodities. It may be argued again fairly that the creditor suffers no injustice. He gets back precisely what he lent, in terms both of money and of goods.

[1] Long-run effects are here had in mind, and especially those long-run effects which are to be expected from steady gains in the efficiency of industry. The proximate effect of increasing money supply, as pointed out in the next section, often has been to cause prices to rise faster than the wages of hired laborers (altho not faster than all money incomes). It is only in the long run that this effect may be counteracted by that of continued improvement in the arts.

He can be said to suffer hardship only in that he fails to share the full advantage of progress. He does not experience, as others do, rising receipts with stationary expenses. The results in the two cases are different; yet in each it may be argued plausibly that the outcome is just, or at least not unjust.

If it be asked which of these two situations—stationary incomes with falling prices, or rising incomes with stationary prices—brings the more equitable adjustment of the relations between debtor and creditor, the answer cannot be given with ready assurance. The problem involves a consideration of the whole question of the right distribution of wealth, and more particularly the question whether equal return for equal labor is the equitable basis for dealings between man and man. In this case, as in many others, we must be content if the outcome is satisfactory on the whole; if clear injustice is avoided, even tho that which is ideally just be not attained.

It is hardly necessary, therefore, to attempt here to resolve this intricate matter any further. As long as movements in general prices proceed slowly, changes of money incomes relative to prices also proceed slowly enough to pass unnoticed by most people. The inverse movement of wages and prices between 1873 and 1896, for example, is perceived only after careful observation of five-year and ten-year periods. The changing relation of money incomes to prices therefore does not commonly entail an injustice serious enough to warrant a departure from the familiar and accepted principles of equity in debt payments. It is in connection with movements in prices that are out of relation to the march of improvement—and particularly the violent movements—that the distressing problems arise. These impose shifts in the distribution of wealth between debtors and creditors that have no equitable basis whatever.

§ 7. It is not to be supposed that the only important effect of rising or falling prices is that on the relations between debtors and creditors. A second consequence is a change in the pace of industrial operations. Fluctuations in the value of money may affect seriously the entire mechanism of production.

Periods of rising prices are commonly periods of industrial ac-

tivity and prosperity. True, the prosperity may be apparent rather than real. People so habitually reckon their incomes and resources in terms of money that they think themselves better off when money incomes go up, disregarding, for a time at least, the fact that their expenses go up also. But it is by no means all a matter of deceptive appearances. The business class feels a stimulus from rising prices; and so long as the management of industry is in the hands of the business class, that which stimulates its members to activity acts as a real stimulus to productive industry. In some degree, no doubt, the effects on business men as on others is psychological. They think they are gaining when prices rise, whether or not they gain as regards the purchasing power of their incomes; and this appearance of gain spurs them to activity. But it is not all a matter of appearance. The stimulus to business men, and hence to production, rests on real and substantial advantages to his class.

These advantages arise partly, tho by no means wholly, from the fact that business concerns are debtors. Business concerns, be it noted, not business men. The position of men as individuals is usually a double one: they are both lenders and borrowers, both investors and applicants for funds seeking investment. In the early stages of a prosperous business man's career the borrowing predominates. In the later stages, if he has prospered, he has accumulated and has come to be a lender and creditor. Probably at no stage, so long as he is active in affairs, does he distinguish sharply between the funds which he has set aside as permanently invested and those which are used for current operations. With the development and ubiquity of the corporate organization of business, the distinction becomes sharper. While the head of a large corporation doubtless has accumulations of his own, and so is a creditor, yet when acting as business manager he is likely to be a borrower for his concern and especially so as regards the funds used in current operation. The corporate concerns, like the individual firms, gain from rising prices because interest charges lag behind prices. Their operations respond to the stimulus and production is quickened and enlarged.

§ 8. In still another way the relation between employers and

employees has effects of its own on changes in prices as well as on production and the march of industry. These effects arise not so much from the fact of fluctuations and oscillations as from inherent characteristics of capitalistic industry.

Consider the various kinds of costs. Wages are costs; just in what sense costs, and how related as costs to other elements which also are reckoned as costs, is a far-reaching question of which the various aspects will present themselves again and again. For the present, we will continue to look at wages as the capitalist producer looks at them,—expenses which he must meet; for him they are like expenses for raw materials, interest, rent, taxes. Now, for the individual producer there is a period of production in the course of which he makes payments for wages and other costs which have to be met some time before receipts come in.[1] The period is longer in some industries than in others. In all, the process of a change in the price level affects and disturbs the relation between prices which constitute costs, notably wages, and the prices of the goods that are sold. If there is a general upward movement of prices in the interval between the initiation and the completion of the period, goods will be sold at prices high in proportion to what those goods cost. If, on the other hand, prices fall in the interval, things are sold at prices which by comparison with their costs are low. In either case, of course, only a part of the cost is incurred at the very beginning or at the original level of prices. The outlay takes place bit by bit, as it were, over the course of time, and each installment tends to be adjusted to the price level just then pre-

[1] This is the "period of production" in the British sense: the interval elapsing between the time of paying out money as costs for the individual concern or industry and that of receiving money from the sale of the goods. It is usually a matter of months. Quite different is the period of production in the Austrian sense: the whole long interval, a matter of years, between the first application of labor in the successive operations of production and the final emergence of a consumer's good. The period in the British sense comes to its close when a commodity emerges which can be sold for cash: a ton of iron or a crop of cotton. The period in the Austrian sense comes to its close (more accurately, near to its close) when a commodity emerges which satisfies a consumer's want: a kitchen stove or a cotton shirt. The first is important for the shorter-period phenomena, such as the immediate relations between employers and employed, and for the present topic,—the oscillations of industry in their relation to price changes. The second bears chiefly on longer-period phenomena, the fundamental questions of wages, profits, interest, which are reserved for later consideration. Vol. II, Chapter 38.

vailing. The fact that production takes place in anticipation of sale and in advance of sale means that if general prices move in the interval there arises a disparity between the prices at which goods are sold and the costs at which they were produced.

This cause does not ordinarily give occasion to very marked divergences between costs and prices. The period of time involved in most processes of production is not very long; seldom is it more than a few months. Movements in the level of prices on the other hand ordinarily take much longer before they impinge markedly on the individual good. In times of great economic disturbance, however, quick shifts of general prices do occur, and bring marked disparities in the cost-price relation even for commodities whose production is of but moderate length. Moreover there are some raw materials like copper, iron, rubber, wool, the prices of which are important and symptomatic costs; and these may change rapidly even while the general movement is slow or may change slowly when the general movement is rapid.

In general, the spread between costs and prices is important so far as the movements of the two occur at unequal pace. As regards some costs, a disparity results from their contractual nature. Technical and professional labor such as that of engineers and supervisors is often engaged at fixed salaries for a considerable time. The contract is binding upon both parties and during its life the salary is a cost that cannot be changed. Capital often is procured in the same way. Funds are borrowed from the public by the issue of bonds or other instruments that involve the payment of a stipulated interest. The interest is a cost fixed for a period, perhaps one of many years. Land and buildings, and sometimes ships and machinery, are hired in like fashion. The rental paid for them is part of the terms of the contract. In all such contractual arrangements the business man—or the business concern—is undertaking to pay, regardless of the ups and downs of prices, a stated sum.

Of greatest influence on the cost-price relation is the movement of wages. In most industries wages constitute the largest single immediate element of cost. That wages go up in times of rising price level more slowly than prices of goods is one of the best

attested facts of economic history. It holds of almost all sorts of hired persons—not only manual laborers, but clerks, overseers, teachers, salaried persons. In the case of technical and professional workers, the sluggishness is due chiefly to the contractual nature of such wages. In the case of clerks, mechanics and other manual laborers it is due partly to the lack of bargaining power among the workers, partly to the force of custom, which is especially strong as to wages. The slow pace at which competition in the labor market works out its results is connected with the many peculiarities in the dealings between employers and employees, and especially with the position of the employer as feeling the brunt of any industrial change. Of the fact here noted there can be no question; when prices rise, the wages of hired workers do not rise as fast.

It is familiar experience that those business men gain most in periods of rising prices whose operations involve in largest proportion the payment of wages. The mere trader or merchant usually gains least; the prices of the things he buys go up almost as fast as the prices of the things he sells. The manufacturer who buys few materials, and whose expenses are chiefly in the direct purchase of labor, profits most of all. Such, for example, is the situation of a highly integrated enterprise like the United States Steel Corporation, which hires laborers directly to dig iron ore, mine coal, convert the coal into coke, transport these materials, smelt and shape the iron and steel.[1] When the prices of the iron and steel go up, the concern gains hugely, since its main outlay, that for wage payments, is relatively stable. Those iron and steel makers, however, who have to buy iron ore, or coal and coke, gain comparatively little; the prices of their materials go up as fast or nearly as fast as the prices of their products. The business man who is nearest the ground, so to speak—nearest the laborer—profits most from the relative stability of wages.

Conversely, in periods of falling prices the business class as a whole commonly loses. Then, since the same forces tend to keep wages stable, a fall in prices brings loss. Probably wages feel the

[1] That is, thru its subsidiary corporations. Between the subsidiary corporations there is nominal purchase of materials.

effect of falling prices less slowly than they do those of rising prices. The employer's superior bargaining power enables him more readily to mitigate the loss, just as it aids him in reaping the gain. But some loss there is, for the same fundamental reason—on him falls the first effect of any change. The two lags, however, do not offset each other. The weak bargaining position of the laborers brings it about that wages, tho they do not fall so fast as prices, tend to fall more quickly than they rise in the converse case.

All these remarks about changes in wages apply more to earlier times,—say to the close of the nineteenth century,—than to later. In the twentieth century the situation changed. For various reasons wages have moved more in accord with prices in times of rise, less so in times of fall. Labor organizations have become stronger, more alert to take advantage of favorable conditions, better able to resist the effects of depression. For one reason or another, the stickiness of wages has become less: markedly less in England as regards periods of falling prices, and less in the United States when prices are rising.

A situation similar in some respects arises at times with regard to another class of costs,—the prices of certain raw materials that have come to be subject to monopolistic control. During the two periods of depression which came after the war of 1914–18 —the first after 1920, the second after 1930—it appeared that the prices of some large classes of materials fell more slowly than did the prices of the more finished goods. The explanation was that in the so-called "heavy" industries there were combinations, agreements, cartels, monopolistic control in greater or less degree, and that these enabled the prices of such things as iron, cement, copper, rubber to be kept up, or made to fall but slowly, in face of a general decline. It is not the mere fact that they are raw materials which sets such articles apart as being peculiarly sluggish in price movements. So far as their prices are immobile the explanation is that they are of the standardized kind and are amenable to great-scale operations; thereby they invite agreements and combinations, and tend toward the cartel or monopoly. Other raw materials, such as cotton or wool, not produced on the grand scale, show no such sluggish movements. If in modern

times these have not infrequently been prevented from falling in times of depression, it was done thru resort to the political power —action by the state in controlling the supply or assisting the holders in controlling it. As regards such political factors, generalization and prediction are even less possible than for the prices of goods produced under monopolistic conditions.

§ 9. It appears then that changes in the value of money are not a simple matter: especially if the term "value of money" is considered, as it should be, to apply to money incomes as well as the prices of commodities. What constitutes stability of money is not the same thing when one thinks of commodities alone as it is when money incomes are included. Nor do the same questions arise for short-period as for long-period movements.

Much has been said, at one time or another, in favor of a "commodity dollar"—i.e. of a monetary unit which shall always command the same quantity of commodities, a "basket" of the same goods. Now it is not at all a matter of course in the transactions between lenders and borrowers that the payment of debts in the same quantity of goods is the one just thing. Nor, even if this ideal be accepted, is it a matter of course that the quantity of commodities *can* be made the same thru the period involved in long-term debts; since the contents of a typical "basket" of goods change as tastes and technique change. Further, when it comes to the changes in prices which concern people most in ordinary times,—those changes which appear over periods of months or of a year or two,—it would seem that irregularity in the movements is inevitable. The rise or fall cannot be the same for each and every commodity, while yet the individual persons are seriously concerned only with the prices of the things which they buy and sell and have little interest in the wide-ranging up or down swells.

On the whole the wise policy, in the world as we have it, is to aim at a monetary system in which abrupt changes over the whole range of transactions are prevented once for all, and great changes in particular articles are prevented from being abrupt, even tho they be of the kind which arise from new and welcome technological improvements. Given these results, it seems to be not a matter of the first consequence whether the general level of com-

modity prices shows in the course of decades a movement, even a considerable one, either upward or downward. Any mechanical or supposedly automatic device, e.g. that for a commodity dollar, is likely to be not only intricate and troublesome in execution but of doubtful service for securing the equity desired.

CHAPTER 21

COIN AND SPECIE

§ 1. Gold and silver, long the dominant forms of money, gradually became of less importance in the nineteenth and twentieth centuries. The qualities that led to their use as money.—§ 2. Their position strengthened by factors other than the physical qualities. Convention and habit.—§ 3. Coinage a public function. Under the gold standard, gold and bullion interchangeable.—§ 4. Bimetallism, and the difficulties to which it led. Its history. Silver suddenly discarded toward the close of the nineteenth century. The attempt to revive it in the U. S. from 1878 to 1934 an anomaly. —§ 5. Some problems of coinage. Subsidiary coin.

§ 1. THIS and the next following chapters will give more detailed consideration to the three salient constituents of the money supply—coin, paper money (notes), deposits. The present chapter will deal with coin.

Thirty years ago, when the first edition of this book was in preparation, coin could still be spoken of as a highly important constituent of any effective medium of exchange, perhaps the most important. After the passage of a most eventful generation, that emphasis is no longer justified. In the English-speaking countries, coin came to be the least important of the three, and the order of weight became the inverse of that formerly given, deposits being now first, paper money second, coin third. Yet coins are by no means negligible and specie held for possible conversion into coin is even less negligible. Not only do they retain a place of importance, but an understanding of the ways in which they were used in former times, of the process by which they have been displaced, and of the ways in which they still play a part, is indispensable for an understanding of the meaning and working of the other constituents and of the monetary structure as a whole.

Historically the chief reason why gold and silver became the money metals was that they satisfied the craving for adornment. Things that minister to the deep-rooted love of display are in unfailing demand. And any commodity that is in unfailing de-

mand may perform passably the functions of a medium of exchange. It is this which explains the wide variety of things that have so served—cattle, grain, salt, furs, tobacco, and what not. The luster and sheen of gold and silver caused them to be highly prized in the early stages of civilization. The glitter of the bauble is the origin of the monetary use of the precious metals, precisely as glass beads and scarlet cloth are serviceable to explorers for barter in those few regions where savagery is still unaffected by the conventional ways of civilized man.

Other qualities contributed greatly to making gold and silver the money metals. They are singularly free from liability to deterioration. Rust does not affect them. They retain their luster with unusual constancy. The durability of the metals was largely responsible for their use as moneys.

Most important of all, however, was the scarcity of the precious metals. They proved to be sufficiently abundant for money use, and yet not so plentiful that they ceased to be prized. Any metal that is fairly scarce might be selected for monetary use. Iron was used in the early days of Rome. Copper was used to a considerable extent in later times; and it is still in use, tho not under conditions that give it much significance. In the course of time, both iron and copper were discovered and produced in such great quantities that they ceased to have any special value from their rarity. Gold and silver remain comparatively scarce. Tho common, and very widely distributed in the earth's crust (gold perhaps most widely distributed of all), they are rarely found in large amounts or under conditions which enable great quantities to be secured at small cost. It is true that highly productive mines have been not infrequently discovered, and in recent times (since 1890) new sources have been exploited to a striking extent. Some of these changes have had far-reaching effects on prices and on the modes of use for the two metals. Some of them, too, have caused the question to be raised, at one time or another, whether silver, or gold, or both, might not become so cheap as no longer to be fit to serve as money. On the whole, however, their scarcity and high cost have continued. Tho now produced in quantities that are enormous compared with those of former centuries, their annual production is

still very small as compared with that of iron, lead, copper, tin, and zinc.[1]

§ 2. The use of gold and silver for money, originally resting on the intrinsic factors of beauty and scarcity, came in time to be largely a matter of convention. Once established as the money metals, they retained their position to a great degree by force of custom. Anything which passes readily from hand to hand has value from its mere acceptability. The strong influence of convention and habit is illustrated by the wampum of the American Indians. These strings of shells, originally sought because fancied for ornament, were in course of time accepted, without thought of their ornamental qualities, as a medium of exchange for the Indian tribes and the early settlers. Among certain African tribes, tiny axes (called bikei) served as the medium of exchange. They were conventionalized survivals from a time when real axes were rare and enormously useful, and that importance had given them the prime quality of general acceptability.[2] Paper money illustrates the same tendency. In the first stages of its use, it had to be really exchangeable on demand for specie; otherwise it would not be taken in payment. But once people were used to it and accustomed to seeing it received by everyone and paid out by everyone, it could circulate as money with little reference to its convertibility into specie. For many centuries specie had the established position which paper money has secured only within very modern times. Merely because all the world accepts a thing as money, it becomes peculiarly fit to serve as money.

Further, the fact that specie serves so universally as money tends

[1] The total production, the world over, of the more familiar metals was in metric tons.

	In 1900	In 1930
Pig iron	41,000,000	80,500,000
Lead	860,000	1,700,000
Copper	486,000	1,600,000
Zinc	471,000	1,400,000
Tin	85,000	196,000
Aluminum	7,800	226,000
Nickel	7,500	54,000
Silver	5,650	7,050
Gold	388	591

[2] See Miss Mary Kingsley's *Travels in West Africa*, p. 320.

to maintain its value by giving it a utility for social prestige. Many of the non-monetary purposes for which gold and silver are used have become of minor importance. Brass and sundry imitations usually do as well. Between the serviceability of plated ware and of solid silver there is no substantial difference. The one great utility which the sterling metal retains is like that of the diamond —it satisfies the love of distinction. The fact that gold and silver are used as money keeps up their value; the fact that they are valuable gives them utility for display; and this in turn helps to sustain their value for monetary as well as for non-monetary uses.

§ 3. When there is a metallic money, or when the monetary system is based on a metallic standard, a fixed relation is established between the currency unit and the standard metal. The most common and important case in modern times has been that of a gold standard, and here the fixed relation is between the currency unit and gold. Where coins of the standard metal are actually minted and used, the relation is established by prescribing in the coinage law the rate at which the mint is required to give coin for bullion. Where the standard metal is not minted into coin, but paper notes are used instead, the arrangement may be different. The gold standard grew up as a system in which the coins were used, and it long retained this form. What the gold standard means is best understood by attending first to this earlier form.

The rate at which coin is given for bullion is the "mint price of gold." In the period when the gold standard was in wide use (roughly from 1873 to 1914), the English mint price of standard gold, i.e., gold eleven-twelfths fine, was £3, 17s., 10½d. per ounce; each ounce was manufactured into sovereigns at this rate. In France the mint price of fine gold was 3447.74 francs per kilogram; in Germany, 2790 marks per kilogram; the figures in both cases indicating how many francs or marks were manufactured from the kilogram of gold. In the United States, the phrase "mint price of gold" did not come into use, because our coinage legislation proceeded not by specifying the number of dollars to be minted out of a given weight (say an ounce) of gold, but by specifying how much gold the dollar should contain. The dollar was required (at this same period) to contain 23.22 grains of fine gold, or 25.8 grains

of gold nine-tenths fine. Dollar pieces proved too small for convenient use, but five-dollar pieces were coined with five times this weight of gold and ten-dollar pieces with ten times the weight. The mint price of gold, if that phrase had been used, would have been $20.67 per ounce.[1]

Because the amount of gold coin given for bullion under the full gold standard never varies (so long as the coinage legislation remains unchanged), people often spoke of the value of gold as unvarying. Accustomed to all exchanges and all values in terms of price, they thought of the value of gold also as the price (the mint price) of gold bullion. But obviously the purchasing power of gold is a very different matter. The value of gold in the economic sense of the term depends on the general range of the prices of commodities; or rather, it *is* the general range of prices. The fact that the mint price of gold is fixed does not mean that the value of gold is stable; all that it means is that the value of gold and the value of money move together. Yet the notion that gold has a fixed and unvarying value, not to be changed or shaken by legislation, was widely and persistently held, and indeed still crops out; being a way of contrasting the supposed stability of gold to a fickleness and instability inherent in paper money.

Free coinage—that is, actual coinage of gold and actual circulation of the coins—tho not essential to the maintenance of a gold standard, is yet of the first importance for its working. When this part of the system is dropped the operation of the rest is greatly affected. Free coinage means that every holder of bullion may bring it to the mint and have it converted into coin without limit as to amount. The cost of manufacturing the coin may be borne by the mint, that is by the public; when so borne, coinage is gratuitous as well as free. Coinage is not gratuitous if the mint returns to the applicant coins containing a slightly less weight of specie than he presented. The difference retained by the mint constitutes a charge toward meeting its expenses. Such a difference or deduction is called a seigniorage (a name derived from the exclusive coinage rights of the king or feudal seigneur). Where a

[1] These figures are illustrative; they indicate the practices of the period when the gold standard was in full bloom. For the post-war developments and the practices and problems of that later period, see below, Chapter 30.

seigniorage is charged, the exchange value of coin may exceed to that extent the value of bullion. So long as this kind of monetary system was in vogue the mints of most countries followed the first-named practice; they returned to the person who presented gold bullion precisely the same weight of fine gold in the shape of coins. Sometimes, indeed, this return was not immediate; there was a delay corresponding to the length of time required for the manufacture of the coin. Thus in the United States a period of six weeks usually elapsed between the delivery of the bullion and the return of coin. Since there is a loss of interest during the period of waiting, such a delay will cause the value of bullion to be slightly less than the value of gold, even tho there be free coinage without seigniorage. These causes of divergence between gold bullion and gold coin—whether seigniorage or delay in coinage—have ceased to be of appreciable importance.

Under the fully developed gold standard, not only can the bullion be converted into coin at the mint without charge or for a trifling charge, but gold coin can be readily converted into gold bullion; either by private melting, or by an arrangement, common at the mints, for giving bullion in exchange for coin at fixed rates. The situation is very different with the subsidiary coins of silver, copper, and nickel, which present problems of their own. As regards gold, where a gold standard is fully established bullion and coin are interchangeable.

§ 4. Thruout most of monetary history, silver was used as a money metal on a footing equal to that of gold. In European countries silver as well as gold was long coined freely at a fixed mint price. Under this kind of arrangement the two metals were interchangeable; both were freely coined, and coins made of either metal were full legal tender. The method in its fully developed form is illustrated by the system adopted by the United States in 1792. The gold dollar was to contain (i.e., if coined, would have contained) 23.22 grains of pure gold, or 25.8 grains of gold 9/10 fine; the silver dollar similarly was to contain 371¼ grains of pure silver, or 412½ grains of silver 9/10 fine. Either gold or silver bullion could be brought to the mint for coinage without limitation as to amount; gold coins and silver coins alike were

established by law as tender for the discharge of debts. This was bimetallism, deliberately adopted and carried out to the full.

For several centuries bimetallism prevailed in more or less changeable and chaotic form. Even after the art of coinage had developed, the actual coins were used in pieces of widely varying weight and fineness and had to be put on the scales at each transaction; and this was the case with silver as well as gold. Gradually the more perfect form developed: both metals were manufactured into well-cut uniform coins of the same or similar names and denominations. Silver was the more plentiful and the more commonly used; gold became plentiful only within the nineteenth century.

However well devised the coinage system was, difficulties continually arose in the endeavor to treat the two metals as homogeneous. These difficulties became accentuated in the nineteenth century, and finally resulted toward the close of that century in the displacement of silver from the position of a freely coined money metal. This change, one of the most notable in monetary history, was brought about in the surprisingly short space of one generation. It deserves some detailed examination because it constitutes a sort of preamble to the quite new situation which emerged in the twentieth century.

When the double standard is adopted, the question arises whether the ratio at which the metals are coined by the mint and are thus given purchasing power in the form of money conforms to their relative values as bullion. If at the mint 16 ounces of silver are coined into as many dollars as one ounce of gold, and if, as bullion, 15 or 15½ ounces of silver can be sold in the market at a price equivalent to one ounce of gold—no one will bring silver to the mint.[1] The silver will be more valuable as bullion than as coin; and experience proves that a very small fraction of difference suffices to decide that the metal shall not be presented for coinage. If, on the other hand, silver can be sold as bullion only at the rate of 16½ or 17 ounces of silver for one ounce of gold, no one will

[1] These figures, selected because of their occurrence in the historical case, are used for illustrative purposes only. The point of principle is the consequence of a divergence between mint and market ratios, and is the same regardless of the specific figures.

bring gold to the mint. The holder of gold can get for it at the mint only as many coined dollars as he can get for 16 ounces of silver. By exchanging his gold in the market for 16½ or 17 ounces of silver bullion he can get more coined dollars; and accordingly he will present only silver bullion at the mint. To repeat, a very small variation between the ratio fixed at the mint and that which rules in the open market will cause one or the other of the two metals to be the sole one presented at the mint for coinage.

The metal which tends under such conditions to be presented at the mint is said to be overvalued. The metal which is not presented, and which indeed is subjected to the opposite process of being melted into bullion from coin, is said to be undervalued. Where silver is coined at a ratio of 16 to 1 of gold, while on the basis of market values 15 of silver is equal to 1 of gold, silver is undervalued at the mint, being given a lower value there than in the market. And where in the market 17 ounces of silver equals 1 of gold, while at the mint silver is still coined at 16 to 1, silver is overvalued.

That metal which is overvalued will tend to become the sole constituent of the metallic circulating medium. It alone will be presented for coinage. This, to be sure, will tend to withdraw it from the bullion market; and this process will tend to raise its value as bullion. Conversely the undervalued metal, not being presented at the mint for coinage, will tend to be more plentiful in the market as bullion; and this will tend to lower in turn its value. By such processes, the market ratio may possibly be brought to equality with that at the mint. But if there be a permanent force at work which brings about a continuing difference, even tho a slight one, then the undervalued metal will go out of circulation, the overvalued metal will come more and more into circulation, and eventually the metallic money will consist of the overvalued alone. The more considerable is the divergence between the mint and market valuations, the more quickly will this process work out.

The tendency of the overvalued metal to drive out the undervalued is often termed Gresham's law. The name is derived from a Sir Thomas Gresham of the sixteenth century who gets thus an

undeserved fame for stating something that had long been known. The "law" is simply the commonplace fact that where mints take two metals for coinage at a ratio differing from that in the open market, the cheaper metal, if there be enough of it, will displace the dearer from monetary use. The cheaper money metal will be used by preference for presentation at the mint; the dearer will be used by preference as bullion for industrial purposes.

The double standard was traditional in Europe for many centuries. It is true that in England, tho the bimetallic system had prevailed thru the eighteenth century, the circulation was composed chiefly of gold; the single gold standard was established by law in 1816. But on the continent of Europe in general the double standard prevailed, with a stock of metallic money made up chiefly of silver. Until the second half of the nineteenth century France alone had a circulation in which gold, tho by no means the larger constituent, was important side by side with silver.

In the United States, when our coinage system was established in 1792, the European model was followed. The complete double standard was adopted and the coinage ratio set at 15 to 1. That ratio was chosen after careful inquiry, but it proved to differ from the market ratio, which was about 15½ to 1. The ratio of 15½ to 1 was accepted about ten years later for the coinage system of France. Silver accordingly was overvalued at the United States mint, and gold was undervalued. No gold was presented for coinage, and the metallic circulating medium consisted wholly of silver. In 1834 the ratio was abruptly changed to 16 to 1. This overvalued gold as much as the old ratio had overvalued silver, so that gold alone was now presented for coinage. Silver gradually drifted out of circulation and out of the country. After the California gold discoveries in 1850, the change became pronounced. Great quantities of gold were coined; silver quite disappeared. In the business world the silver standard was replaced by the gold standard. The *de facto* situation was recognized in legislation when arrangements were made (in 1853) for the use of silver as a subsidiary coin.

In 1873, toward the close of the paper money régime of the Civil War, the coinage legislation of the United States was over-

hauled and consolidated. In this revision of the statutes, the silver dollar was dropped from the list of coins that could be struck. Therewith bimetallism was formally ended by law. A few years later, after strong agitation for renewed use of silver, great purchases of that metal were made by the government, and coins were again struck which had full legal tender quality; but free coinage was not restored. Under two legislative measures, one of 1878 and the other of 1890, no less than 570,000,000 dollars of silver money were injected into the circulating medium before the process was stopped in 1893. Actual silver dollars were little used and comparatively few were coined; the silver circulated chiefly in the form of silver certificates, and might be regarded as a hybrid kind of money,—a cross between subsidiary coin and full legal tender. Tho the amount in circulation was large, it served in good measure merely to replace national bank notes which were declining, and the effect in the monetary structure was thus largely neutralized.

In 1934, in the course of the monetary upheaval of the early years of the first Roosevelt administration, a new and anomalous step was taken. It may not unfairly be described as atavistic,—reversion to an outlived type. The advocacy of a return to bimetallism proper, which had played so large a part in the monetary controversies of the nineteenth century, had quite ceased; the world had settled down to the use of silver in the Western countries for subsidiary coin only. Legislation on money had turned into quite different channels. With hardly a word of debate or explanation, an act of 1934 directed the Treasury to purchase silver in the market until enough was accumulated to bring the quantity of silver thus held in monetary reserves up to one-quarter of the total "value" of all the metallic reserves.[1] Silver certificates were issued on the basis of the metal, retaining their hybrid character.

[1] There was an alternative, very unlikely of realization: the purchases were also to cease if the market price of silver should reach the figure $1.29. The entire procedure was tortuous. The silver was purchased by the Treasury at the market price; but once in the vaults of the Treasury, it was valued for bookkeeping purposes at $1.29 per ounce. It is at this arbitrary "value"—much above the market price—that the amount of silver held by the Treasury was calculated in measuring the size of silver reserves. Hence, the Treasury was required to buy much less silver than if the Treasury used the market price for the calculation.

The silver legislation of the earlier years, from 1878 to 1890, is explicable on historical and political grounds of a kind hardly to be found anywhere outside the United States,—the demand from pioneer regions for more money, selfish pressure from the silver mining constituencies, a curious sentimental clinging to silver as a political idol, successive compromises. The legislation of 1934 rested on a similar jumble of forces. It was based on no defensible policy, and could lead to no end but further huge accumulations of useless metal.

France suspended free coinage of silver in 1873, the same year as the United States; the coincidence in time was fortuitous. She had maintained the double standard in 1803 when she established her present system of decimal coins. The coinage ratio was then set at 15½ of silver to 1 of gold. Until 1850 the circulation was in fact chiefly of silver, because the price of silver in the market tended to be somewhat lower than the equivalent of the coinage ratio. After 1850, however, the situation changed. The unexampled supplies of new gold from California and Australia were poured into the world's markets. So strongly was the bullion market affected that it became advantageous to send into France for coinage not silver but gold. In France, just as in the United States, gold began to displace silver and the displacement was accepted with the same equanimity. After a decade or two gold came to be regarded thruout the world, reasonably or unreasonably, not only as the preferable metal, but as *the* metal on which to base a good monetary system. The German Empire, when reorganizing its currency system in 1871, adopted the gold standard outright, influenced chiefly by the English example. The United States had been since 1850 practically on a gold basis. International transactions between all countries were conducted in terms of gold.

Hence when great discoveries of silver were made, and production increased after 1870 at a striking pace, the threatened inversion proved unwelcome. Silver began to flow into France in great quantities; gold began to flow out. France, faced with a distasteful prospect, dropped the free coinage of silver. Other European nations followed her lead. Belgium, Switzerland, Italy, and Greece had been joined with France in the Latin Union, the

object of which was the adoption of a uniform decimal coinage system based on the French franc. The Union had kept to bimetallism at the ratio of 15½ to 1 and the several mints acted in common. When France suspended free coinage of silver, therefore, the others followed suit. Bimetallism came to an end.

The last important event in the chain which deposed silver came in 1893 when British India put an end to the free coinage of silver. Continued free coinage, with the steady fall in the price of silver, had caused so large a volume of silver to be presented at the Indian mint as to result in serious embarrassment. The mint was closed to silver.

Thus ended the rapid transformation. Silver, for centuries a standard money metal and long the predominant metal in the circulating medium, had lost its place. The fundamental cause of the change lay in the conditions of production of gold and of silver. It was the increased output of gold after the great discoveries of the middle of the century that brought about the displacement of silver from actual circulation; it was the prospect of a similar vast increase in silver production that impelled the final demonetization of silver. Before 1870 the annual supply of that metal from the mines had been about 30,000,000 ounces. After 1870, it rose thus:

	Million Ounces
Average annual product in the 5-year period 1871–1875	63
" " " " " " " 1876–1880	79
" " " " " " " 1881–1885	92
" " " " " " " 1886–1890	109
" " " " " " " 1891–1895	158
" " " " " " " 1896–1900	165
" " " " " " " 1901–1905	168

So great a fresh amount pressing on the market, with most mints closed to free coinage, caused a steady decline in price, and led to a growing practice, indeed a policy which seemed firmly established, to restrict the use of silver for monetary purposes to the field of subsidiary coin.

§ 5. It remains to say a few words about some coinage problems, and about subsidiary coin.

The tendency for overvalued metal to drive out the undervalued may appear, tho in a rather different way, even where only one metal is coined. When the metal is coined at its actual value, that is when the value of metallic coin as money is no greater—or is intended to be no greater—than its value as bullion, light-weight or abraded coins tend to displace full-weight ones. Until the nineteenth century the machinery for manufacturing coins worked slowly and somewhat imperfectly. It was difficult to turn out a great many coins rapidly; and the coins minted not only were subject to ordinary abrasion, but in consequence of uneven mintage were subject to clipping. New and good coins were therefore likely to be picked out for use in the arts or for exportation, while only the poorer pieces remained in circulation. In the United States—to give an example—during the period from 1792 to 1834, when silver was the only money metal in actual circulation, foreign silver pieces of various mintage were in actual use. Because at the beginning no United States mint or coins existed, these foreign coins had been authorized for use in public payments. When the mint was established, and coins were issued from it, the new coins could not displace the foreign pieces, being full weight and preferably used for the arts or exportation. Hence the coinage seemed futile and was discontinued, only the more or less imperfect foreign coins remaining in circulation. Difficulties of a similar sort were long encountered in all European countries, from the Middle Ages thru the eighteenth century.

The remedies for these difficulties are simple: first, the plentiful and accurate manufacture of full-weight coin; second, the withdrawal of all legal sanction (such as receipt in payment of public dues) from worn coin; and third, the redemption at no charge of pieces which become abraded by ordinary wear. It was once common to enact that pieces which had suffered in weight beyond a certain tolerance should not only lose their validity as legal tender, but should be accepted at the mint only as bullion, not at their face value. This, however, served to perpetuate their circulation. The holder, called upon thus to suffer the loss in value from abrasion, tried to pass them on to other persons; and others accepted them simply because they were ready money. The better

practice, therefore, is for governments to redeem at face value all coins which have not been intentionally clipped or sweated. The common adoption of this practice, together with the ample adequacy of machinery for providing new and uniformly good coins, has removed the difficulties here described.

In recent years the metallic circulation has come to be almost exclusively made up of subsidiary coins. Most nations have given up the coinage of the standard metal (gold), using it only for reserve holdings. Subsidiary coins—silver, nickel and copper—came into use because the high value of gold makes it unavailable in minor payments; as indeed had been the case with silver in earlier times. Gold pieces worth less than five dollars are very small; they are too easily lost and too subject to wear to be thoroly satisfactory. Yet a multitude of transactions must be settled with money of smaller denominations. Silver is convenient for such money, hence continues to be used by most civilized countries as a minor coin. For the smallest transactions even silver has not bulk enough; for these resort is had to nickel and copper.

The subsidiary system of the United States was created by an act passed in 1853. Silver coins were authorized—half-dollars, quarters, and dimes—containing so light a content of fine silver that no one would be tempted to export them or to melt them for industrial use. The silver half-dollar, for instance, was made to contain (and still contains) 172.8 grains of fine silver, or 345.6 grains for two half-dollars. The silver dollar, whose free coinage at that time was still authorized, contained (and still contains) 371¼ grains. Thus the silver subsidiary coins were made worth considerably more as money than as bullion, because the coins contained considerably less than their face value of actual metal. This situation has continued ever since; indeed the overvaluation is greater now than it was at first because the price of silver has continued to fall.

Obviously, the government makes a profit by an operation of this sort. The overvalued silver coins are paid out by the government in its ordinary disbursements, or are exchanged by it for full-value gold. In either case there is a profit. This also is often called a "seigniorage," tho it differs in important respects from the

seigniorage which may be charged on the freely coined and full-value pieces, when such exist.

If all silver coins were minted freely at the artificial high rate, silver would be overvalued, and would displace gold or even "full-value" silver. But something very different from free coinage operates in the case of subsidiary money. No private person is entitled to present silver at the mint for conversion into small coin. The government itself buys the silver bullion in the market at the prevailing price, and alone arranges for its coinage. The amount which the government thus buys and coins is limited to the quantity found necessary to meet the needs of small-change transactions. Further to guard against possible abuse, it is provided that subsidiary coin should be legal tender only up to a limited sum, now fixed at $10. Thus there is no tendency for the silver money to displace gold.

Such are the essential principles of subsidiary coinage. Substantially the same system has long been followed as regards the copper and nickel coins adopted for petty transactions. These have been token coins ever since gold and silver came to be used as the standard metals. In fact, the underlying principle—an artificial value due to limitation of quantity—has been followed by all nations ever since subsidiary coin came into use.

The regulation of this kind of money is carried on with variations of detail in different countries. The quantity coined is sometimes fixed at so much per head of population. Thus in Germany subsidiary silver was minted in 1910 at the rate of 15 marks (formerly 10 marks) per head of population; in France at the rate of 7 francs (formerly 6 francs) per head. In Great Britain no specific limit is set; the Bank of England arranges for the coinage of such amounts as experience from time to time shows to be needed. In the United States also no limit is set.

To prevent any possible depreciation of the subsidiary coin in terms of standard money, it is usually redeemable at its face value by the government treasuries when presented in reasonable amounts. In the United States, for example, subsidiary silver coins are redeemable in sums of $20. The same object is achieved by receiving them without limit in payment of public dues, as in France.

CHAPTER 22

BANKING OPERATIONS. BANK NOTES

§ 1. Functions of banks. Savings banks typical as channels of investment. Other similar banking agencies.—§ 2. Commercial banks; their customers and depositors. Their demand liabilities (bank notes and deposits) constitute part of the effective money supply.—§ 3. Notes of private banks, long important, have virtually ceased.—§ 4. Bank notes. Their ready circulation, and their reckless issue during the first part of the nineteenth century, gradually led to regulation and control.—§ 5. Complete security for holders of bank notes finally achieved in the United States. Should depositors have the same security?

§ 1. BANKS perform two functions, equally important, yet very different. They act as agencies for the collection of savings and for investment; they also provide a part of the medium of exchange. The two functions are often performed by the same institution, but not infrequently are separated. A savings bank has to do primarily with investment; and this is the case with many banks of continental Europe which do commercial business also. A strictly commercial bank is not primarily concerned with the sort of "investment" to which the term is commonly applied—that which looks to the creation of permanent plant.

Before turning to the monetary aspects of banking, something may be said about the investment operations. A savings bank accepts deposits; that is, it receives sums of money and promises to repay them. The promise is usually subject to conditions, as for example that the bank reserves the right of requiring notice (ten days, or some such period). Such a deposit is not part of the supply of money, not being subject to transfer by check. It is not expected either by the banker or the depositor that the latter will wish to have his money back promptly. Ordinarily he leaves it with the bank for a considerable time, and expects to get interest on what he has deposited. The operation is typical of the process by which the saving or "putting aside" of money leads to the creation of capital. The money deposited is not kept by the bank, but is lent

to persons who mean to use it in "real" investment, as in erecting factories, warehouses, dwellings. It goes into circulation again and repeats its rounds in performing the functions of the medium of exchange. Meanwhile it has been the instrument by which some persons, having procured the command of the purchasing power which savers thus turn over to them, were enabled to add to the substantive capital of the community. Often the proximate use by the savings bank is in the purchase of securities; that is, of promises to pay, or certificates of ownership, which have been issued by still other persons. In this case the bank is but one link in the chain which connects the savers of money with the makers of capital. The same process of collecting surplus means and attending to their investment is carried on by government postal savings banks, except that here the money deposited is commonly used in buying government securities, and the effect in adding to the real capital of the community—its apparatus of production—is more doubtful.[1]

Many banking institutions other than savings banks, both public and private, engage in operations of this kind. The bankers and brokers who deal in so-called investment securities act as middlemen for the well-to-do, just as the savings banks do for persons of smaller means. The so-called mortgage banks of continental Europe, organized as public or semi-public institutions, perform the same function. Many of the great corporate banks of Germany, France, and Austria long conducted extensive operations in placing investments. They accepted deposits in sums large or small, and either sold securities directly to the investor, or undertook to pay him a stipulated rate of interest. The great historic banking houses of England and the United States, and of the Continent also—the Barings, the Morgans, the Rothschilds, and their numberless rivals and associates—carried on chiefly investment operations. During the greater part of the nineteenth century they were largely concerned with public loans, buying from both strong and weak governments great blocks of bonds, borrowing heavily on their own account in order to be able to put funds at once at the disposal of the needy governments, and re-selling the bonds as quickly

[1] See Chapter 39, Vol. II, Interest on Capital Further Considered.

as possible to private investors. In the course of the nineteenth century, and more and more in the twentieth, they extended their operations into the promoting field,—the nursing and supporting of new industrial enterprises. In the early stages of such enterprises, they make advances from their own means to the active promoters, or (quite as often) borrow from the commercial banks and then re-lend to these promoters. The distinction between their operation as lenders and as promoters usually becomes shadowy; they tend to be more or less active promoters and managers themselves. In time, but very likely after a considerable interval, if all has gone well they sell to the saving and investing public a large part of the securities based on the enterprises they have fostered. Each banking house of this sort usually has its circle of customers and friends, who have faith in its judgment and honesty, and are guided by its advice.

§ 2. Different is the nature of the operations ordinarily carried on by the commercial banks. Typically these do not lend for permanent investment. If they lend to manufacturers, it is commonly not by way of aid in erecting plant but for convenience in current operations. They make loans for enabling raw materials to be bought, the loans to be repaid on the sale of the finished goods after a few months. Similarly they make loans for paying wages and like current expenses, again with the understanding and expectation that repayment will be made shortly. They lend to wholesale merchants who buy large quantities of raw materials to be sold to manufacturers or buy finished goods to be sold to retailers. On the other hand, they serve as repositors of spare funds for people of all sorts, both business men and those not in business. The business men keep a balance in the bank which is larger or smaller according to the momentary stage of their trade; likely to fluctuate considerably, but in each case likely on the average to be substantial. Salaried and professional people and those of independent means maintain a bank balance for convenience, also larger and smaller according to their momentary position but not as much subject to rapid fluctuation as the business deposits. Commercial banks are found in all large trading centers. In earlier days seaports like London, Liverpool, New York, New

Orleans, Boston, each had its set of large strong banks whose doings were mostly of this kind. They have by no means ceased to exist, but they rarely have remained true to type. The causes and consequences of their modern complexity will be considered shortly, bearing as they do on the working of the deposit system and the distinctive banking problems of the twentieth century.

It is the operations of banks of the commercial type which are most familiarly associated with the problems of money—banks supposed not to enter the field of permanent investment, but chiefly to make loans on short time to established business concerns. Such banks receive deposits, but deposits of quite a different character from those of savings banks. They are subject to check, can thus be transferred readily from person to person, and constitute part of the effective money supply. In some cases, also, such banks issue notes, which are more obviously a part of the money supply.

It is with the monetary phases of commercial banking that we now have to deal. They are two. First, notes of commercial banks, in earlier times the more important and still in many ways instructive. Second, now much the most important, deposit money. It is the first of them which will be dealt with in the rest of the present chapter.

§ 3. Note issue is the simplest operation. As carried on by banks of the commercial type it played a large part in the monetary history of Great Britain and the United States during the first half of the nineteenth century. So far as private banking houses or corporations are concerned it has virtually ceased. But for many years such banks put notes out with consequences so serious for the public at large that they loomed up largest among the things that must be regulated and restrained, while the other operation of the same kind—deposits—received comparatively little attention. The history of private note issues is instructive because it shows the same complications and problems that arose later in the case of deposits. Essentially similar effects on the monetary structure and on monetary circulation arose from the notes as later from the deposits, and public regulation set in step by step on lines foreshadowing what was to appear in the later

and more complicated case. At the same time, traditions left over from the earlier stage, and the large place which notes had in discussion and legislation, caused an undue emphasis on the regulation of notes, which persisted after other parts of the monetary system had so grown that the emphasis needed to be placed on the other side, the deposits—bank notes having become not indeed unimportant yet by comparison a minor matter. In this case, as in so many others, economic development proceeded ahead of our understanding of its consequences and hence far ahead of the legislative measures it called for.

§ 4. A bank note is a promise to pay a specified sum to the bearer on demand. In law it is like any other promissory note payable to bearer. Title to it passes in full by delivery, and each successive holder acquires the same rights against the bank. If the institution which issues it is well known, the note may pass from hand to hand for an indefinite time, performing continuously the essential functions of money. Even if the institution is not well known, the note may remain long in circulation if people have become accustomed to the use of such paper substitutes and if there is no special ground for distrusting the particular bank that issues it. Money being to so great a degree a matter of custom, that which one person offers in payment and the next person is likely to accept in payment passes readily from hand to hand. Experience amply proved that not only notes issued by responsible institutions but those issued by others that assumed the outward show of responsibility passed into the channels of circulation with surprising ease.

The bank, none the less, is under a strict legal obligation to pay every note, whenever presented, in that money which is legal tender for debts in general. Assume specie, or gold, to be the only legal tender (the case here is the same if government notes or central bank notes are the legal tender money). The bank must keep at all times some gold wherewith to pay (or, as it is said, redeem) any notes of its own issue that are presented. If it keeps as much specie as it has notes outstanding, the note issue obviously can be no source of profit; there is no compensation for the expense of printing the notes and of maintaining the bank office.

But if it keeps less than the notes outstanding, there is the chance of profit. The excess of notes issued over and above the specie kept on hand is often called the "uncovered" issue. The larger the uncovered issue, the greater the opportunity for gain. Every bank which is left to go ahead without legislative restriction tends to keep as little specie (or other legal tender) as possible and to have as large an uncovered issue as possible.

Bank notes take the place of specie very much as does inconvertible paper (to be presently considered). An extreme case may even be imagined in which they would entirely displace specie. That extreme can never be reached, so long as the banks are held to their obligation to pay on demand; some legal tender money must always be kept. But where banks are allowed to issue without restraint, a near approach to the extreme may be reached. So it was in the United States before 1860, when a multitude of banks were chartered by the several states and issued notes, and each was under the temptation to put out its notes as freely as possible. The everyday circulating medium consisted of these notes, only a narrow margin of specie (in those days the only legal tender) being held in the bank vaults. In some parts of the country, especially in what were then the new regions of the West (Illinois and Wisconsin, for example), redemption of the notes in specie was not insisted on by law and business custom, and the notes were virtually inconvertible paper. In New England, New York, and the eastern seaboard generally, and in Ohio and Indiana, the notes as a rule were really convertible into specie; yet their specie basis was small as compared with all the demand obligations of the banks.

Devices of one kind or another were resorted to from time to time in order to check the reckless issue of notes. One was the Suffolk Bank system of New England, much referred to in the earlier literature on the subject and still interesting and instructive.[1] Another was a statutory requirement (e.g. in Massachusetts) that no bank should pay out over its counter the notes of any other bank. That requirement brings it about, for a banking system where deposits as well as notes are largely used, that the

[1] See Dunbar's *History and Theory of Banking*, 3d ed., Chapter 5.

banks in which notes of other banks are deposited send them (thru the clearing house) to the issuing banks; thereby these are prevented from keeping in circulation more than are wanted by their own particular group of customers. Still another was the prohibition of notes of small denominations. In England, for example, no notes smaller than £5 were allowed; an effective device for keeping sovereigns and subsidiary silver in steady circulation and use.[1]

It may be noted, by way of digression, that the last-named device—the prohibition of small denominations—was the common practice (or legislative requirement) in European governments having public or quasi-public central banks. In the days when the gold standard was in effect, Bank of England notes could not be issued under five pounds, Bank of France notes under fifty francs, German Reichsbank notes under twenty marks. Where this was done, and no other form of paper money was in small denominations, a large circulation of gold coin and subsidiary silver was assured.

§ 5. In Great Britain and the United States note issues by private banks both large and small—that is, by banks other than the Bank of England and the two Banks of the United States—played a large part during the first half of the nineteenth century. They still play a large part in Canada.[2] Where such notes came to be in familiar use, no one could refrain from accepting them. To do so would mean inconvenience not only to him who offered them but also to him who refused, and quite possibly even loss; no other kind of money was to be had. Almost inevitably the notes circulated freely. Thereby successive holders became, one after the other, creditors of the issuing bank. Yet few were aware that this was their legal position. If they had any uneasiness about the goodness of the notes, they merely passed them as quickly as possible to some one else. The notes went round and round. Then, if anything happened to the bank—if it suspended payments temporarily

[1] Compare what is said on this subject, infra, in Chapter 30, on the Gold Standard.

[2] On the continent of Europe the general traditions of control over industry and more especially over money prevented any larger development of this kind. Private banks of issue there were, but were superseded at a comparatively early date by banks owned or controlled by government and alone allowed to issue notes.

or failed for good—he who had a note in hand found himself involuntarily and unwittingly in the position of creditor to a dubious or bankrupt debtor. Such was his position as a lawyer would explain it. The common man merely found that his money had gone bad, without warning and without fault or negligence on his part.

Experiences of this kind were common in the period of private and bank notes. They were so in England and Scotland, still more so in the United States. So great was the distress often entailed, so obvious the fact that the distress came thru no fault or negligence of the note holders, so serious the evils for industry and trade, that legislation gradually was enacted to control or abolish the issues. The English Bank Act of 1844 prohibited any new private notes, and provided that as private banks ceased to exist —were wound up or merged with other banks—the amounts of their notes outstanding (to be exact, two-thirds of that amount) should become issuable again by the Bank of England only. In the United States the evils had become widespread and finally unendurable. The first legislative step for mitigating them directly was the requirement by the State of New York, before the Civil War, of a special fund for the payment of notes, to be put in the hands of State officials and held by them for the purpose. The same method was made part of the national bank system established during the Civil War. The national banks were required to put United States bonds into the hands of the Comptroller of the Currency at Washington, to be held by him as security for the payment of banks that failed. Thereby the bank notes were made as good as legal tender money,—whatever that might be. At the same time a tax was imposed on notes of other banks (i.e. such as were organized under State laws) so high as virtually to prohibit their issue. Tho the national banks were thousands in number, many of them small and by no means immune from bad management and failure, the note holders never could lose. This monetary evil, almost scandal, was wiped out.

It is to be borne in mind that the security thus given was to note holders only. The creditors of a bank are in the main two: note holders and depositors. The note holders were given a privi-

leged position,—a prior lien on good assets of the banks. The depositors were left to themselves. This discrimination was not the result of any deliberate theory. Doubtless there was a hazy consciousness that note holders came to be creditors thru a process of whose meaning they were not aware; perhaps some feeling too that many of them were comparatively poor, or quite innocent of any understanding of business matters. There was perhaps a similar hazy feeling that depositors were in a different position; that they come to be creditors of a particular bank by some deliberate act of their own, and in general are more business-like. But in the main the discrimination arose because it was not perceived that the two classes of creditors were in the same position; that both had claims on the bank payable on demand; and that the monetary system was affected by both sets of claims in essentially the same way. The special position of notes in banking legislation arose because the working and importance of deposits, tho clearly pointed out by various thinkers and writers, was outside the ken of even the most intelligent of the legislators. Only by slow degrees did the deposit system come to be understood. It was only by slow and faltering steps that notes early came to be regulated and safeguarded by legislation and that this was achieved at a much later stage for deposits.

In the next following chapters we shall proceed to a consideration of bank deposits. Before doing so, by way of warning and partly of anticipation of what is to come in still later chapters, a word may be said about the most recent changes, especially in the United States. Bank notes have come in this the twentieth century to have a new place in the circulating medium. Virtually they have ceased to be issued by banks which are under obligation to pay them on demand in legal tender money; that is, under the gold standard, to pay them in gold. They have come to be issued by public banks—public even tho they retain the externals of a private corporation; and commonly they are themselves a legal tender. But a government is a law to itself; no one can compel it to keep promises. The quantity of notes, moreover, may be greater or less according to the needs or wishes of the government at the time. They are no longer a sort of subsidiary money, the amount of

which is settled by the public demand for minor currency. The denominations in which they are printed are indeed determined by the public wishes to have pieces of one size or another; but the total volume is settled by the government. All this is true in the United States at the time of writing (1939). At the same time it is not the whole truth; for the intricacies of our monetary system are great and beyond precedent. The system is not only complicated, but it is still in process of development. On these aspects of the situation more will be said later.[1]

[1] See what is said in detail in the chapters on Central Banks and the Federal Reserve System, on the Theory of Prices, and on the Gold Standard.

CHAPTER 23

DEPOSIT BANKING AND DEPOSIT MONEY

§ 1. Deposits may be lodged, the original form; or they may be created. Continuing deposits and their circulation.—§ 2. How long a created deposit remains effective as money. The relation between created and lodged deposits.—§ 3. Is the creation of deposits semifraudulent?—§ 4. Overdrafts in England.—§ 5. Clearing houses and the interrelations of cash and deposits.—§ 6. Bank notes and government paper of the smaller denominations now play in English-speaking countries a part similar to that of subsidiary coin.—§ 7. Commercial banks and permanent investment. Command of capital precedes savings. Investment is promoted and hurried, with consequences that reach far, and not seldom are injurious.—§ 8. Changes in the character of banking and in the characteristics of bankers.

§ 1. We turn now to deposits, or what is sometimes (and appropriately) called "check money." Under the practices that prevailed until the middle of the nineteenth century, there was a fundamental resemblance between notes and deposits. The resemblance, tho long noted by careful observers, was little understood by many writers on banking. Notes were early the subject of special legislation and regulation; whereas deposits were left to go their own way. Deposits were taken in hand much later. By a curious turn in monetary development, note issues of all kinds, bank notes included, were gradually taken over by governments, and thereafter deposits became the subject of legislation and regulation.

Most persons think of a deposit as cash left with a bank. This the word signifies; and this the transaction originally was. Historically, deposits began as specie left with trusted persons. This was the case with the bankers of Venice and Florence in the early period of the Renaissance, the great bank of Amsterdam, and the goldsmiths of London during the second half of the seventeenth century. Where the banker or goldsmith kept intact the specie so left, he evidently made no gain; nay, he would probably demand a fee from the depositor for the care of the gold or silver. The next

stage came when a depositor, having a payment to make, found it convenient to give to the payee an order on the banker, or to deliver to him the banker's receipt. It was an equally natural process for this third person, if he had no immediate need of the money, to continue to leave it in the banker's charge, simply getting another receipt or having his name inscribed instead of his debtor's on the banker's records as a depositor. If many persons did this, having faith in the banker's honesty and solidity, the banker might use part of the specie in his own ventures or lend it out to others. In the earliest times, the persons to whom such deposits were intrusted were commonly engaged in active business and used the funds in their current operations. Later, they developed the safer practice of merely lending the funds, on short time and on good security. Only as they became systematic lenders did they come to be bankers in the modern sense. Specie was then kept on hand merely in such quantities as were supposed necessary to meet the demands of persons actually calling for it; and the deposits became a source of profit.

This sort of depositing still plays a considerable part in contemporary banking operations. In the United States and England, every person of the property-owning class keeps a bank account merely for the convenience of not handling and safeguarding large sums of cash. Salaried persons and those of the leisure class who have considerable means take to the bank and deposit at once whatever money or rights to money may come into their hands, making most payments by checks on the bank and drawing cash only for petty expenses. They habitually leave most of their current funds on deposit. The banker knows by experience that only a certain fraction (and a surprisingly small fraction) will be called for at any one time. A very great part of what is deposited can be lent out again by him for profit.

But the larger part of the deposits in the commercial banks of a country like the United States or England do not arise in this way. The deposits are in the main *created* by these banks.

It is easy to see in what manner bank notes are "created." A bank's main business is to lend, and to lend not its money or its capital but its credit. This is what it does when it puts out bank

notes. It is usual to say that a bank "issues" its notes. In fact, it turns over to the borrower its own promises to pay, and these promises circulate because the credit of the bank is good. The bank lends him, in other words, *its* credit, which serves the borrower or customer as well as money.

Essentially the same thing is done when a bank lends in the form of a deposit. The common and typical operation is that of the discount of a note. The borrower brings to the bank his promissory note, perhaps signed only by himself, perhaps fortified by the indorsement (i.e. guarantee) of others. The bank then credits him with a "deposit" of the amount of his note, less the agreed interest.[1] He has the same right to draw on the bank as if he had actually deposited money. That right he may exercise either by demanding cash at its counters, or (more probably) thru a check, which directs the bank to make payment to others. The first step in the ordinary commercial loan is the creation of such a depositor's relation with the bank.

It is obvious that this first step will have no special consequence if the depositor exercises his right at once. If he draws out immediately the full amount credited to him, the result is the same as if he had carried off cash without the intermediate step. And it may appear that this is what he is likely to do; for he borrows with the purpose of using the money in business operations. But any depositor who did this, and who had no other relation with the bank, would be an unprofitable customer, and not one to whom the bank would habitually extend "accommodation." All banks, and especially the commercial banks of deposit, deal chiefly with their own circle of customers. These are both borrowers and depositors, both creditors and debtors. They keep their accounts with the bank, and there is a tacit understanding, not infrequently an explicit bargain, that the amount of loan accommodation extended to them shall be in proportion to the balance which is on the average to their credit as depositors.

[1] The interest in case of bank discount is usually calculated on the face of the note, not on the amount lent or credited. Thus if a note for $1000 is discounted for three months at 6 per cent, the interest (1½ per cent for the quarter year) is calculated on the $1000, and the depositor is credited with $985. When the discount proceeds in this way the borrower in fact pays a slightly higher rate of interest than the stated discount rate of 6 per cent on the amount put to his credit.

It is possible, even probable, that very soon after a loan is made the borrower will draw heavily against it. He is not likely to draw out the full amount; for every man, and especially every business man, wishes to keep some balance at the banks as a reserve for contingencies. Even if he draws out a very large part, his bank balance does not long remain depleted. Payments to him from his customers and debtors flow in from day to day, and are deposited in the banks as they come in. Meanwhile, as the days and weeks pass, he must prepare for the maturity of the note with which the transaction began. He does so by accumulating deposits, and toward the end of the period during which the note runs he has larger amounts to his credit. When his note becomes due, he pays it by drawing against the accumulated deposits; that is, he offsets the debt which he owes on his note against the debt which the bank owes him on deposit account. Therewith the transaction is wound up.

But this transaction does not stand alone, and this business man does not stand alone. He will resort to the bank again for loans, and others will also resort to it; for all men in active business are borrowers, in order to carry on their operations continuously and on a larger scale than their own means permit. Their transactions with the banks are repeated in an endless series. Suppose now that a number of such persons are dealing with a bank as borrowers and depositors. While some are discounting and are drawing heavily on the deposits created for them, others are preparing to meet their maturing notes and are depositing heavily. Some happen to have made large payments in the ordinary course of business, and their deposits are scant; others have been receiving large payments, and their deposits are heavy. At any given time, the bank has a volume of deposits large or small according to the business it has built up, and has corresponding resources in the way of notes discounted. Probably it has also some deposits of the non-business kind, independent of its lending operations; and probably it has also made some loans not related to its deposits. But it has continuously a volume of resources (debts to it) closely related to a corresponding volume of deposits (debts due by it).

These continuing deposits are like money. They are essentially

like bank notes, and they serve as part of the medium of exchange just as any other circulating medium does. It may seem odd to speak of a deposit as part of the circulating medium. Most persons would accede to the statement that a check serves to effect payments as well as does a gold coin or a paper note; but they would say that the check, not the deposit, is the equivalent of money. Yet a moment's reflection will show that the check bears the same relation to the deposit as the coin used in making payments bears to coin carried in the pocket. Not all the coin (taking coin as typical of the cash that passes from hand to hand by delivery) is buying commodities all the time. Part of it is carried in pockets or kept in tills, by way of reserve, to be used at convenience. The portion actually used in purchases is determined by what we have called the rapidity of circulation of the money. Deposits similarly are potential means of payment, drawn upon at convenience. Just as in reckoning the total quantity of specie in a community we count the whole supply on hand, not merely that which happens to be making purchases at a given moment, so in reckoning this form of the circulating medium we must count up the total volume of deposits, not that part which happens to be in immediate use in the form of checks. The check is simply the deposit in actual use, and the proportion of checks to deposits represents the rapidity of circulation of deposits.

Rapidity of circulation is high in the case of commercial banks and business men's deposits. Checks are drawn against such deposits daily, and fresh deposits are made daily. In the language of the commercial world, these are "active" accounts; their turnover is rapid. The deposits of persons of the leisure class are much less active. Spare cash or pocket money—whether coin or bank notes or government paper—probably has in all cases a much less rapidity of circulation than a commercial bank's deposits.

§ 2. How long does any particular deposit, once it has been created, continue to be an effective part of the circulating medium? At first blush one would be disposed to say: only until a check is drawn against it. Not so. For the monetary structure as a whole, a deposit does not cease to exist because of the mere fact that a check is drawn. It disappears for good only if a check is

drawn *in settlement of a debt* due to the bank which receives the check. The representative case of this kind is where a customer who has borrowed from a bank pays his debt by a check drawn in favor of the bank itself; or, what amounts to the same thing, when the bank debits his account with the sum due. The payment to the bank may be made in discharge of any sort of obligation. It may be for securities bought from the bank; then too the payment is made in discharge of an immediate liability. A deposit ceases to exist when a debt to the bank is disposed of by its cancellation.[1]

Different is the effect of a check drawn by the depositor in the ordinary conduct of his affairs. There are two possibilities. First, he may cash a check. The typical case is that of pay-roll money. A manufacturer employing many persons draws on his bank every week or fortnight or month for cash wherewith to pay his employees. Then so much of his own deposit is indeed cancelled. But its place is taken in the circulating medium by just as much hand-to-hand cash. That cash finds its way into the hands of retail dealers or of wage-receivers, and either remains in circulation or is again deposited in a bank. It then replaces the previous deposit; originally a created deposit, it now appears as a lodged deposit. Transmutations of this kind are going on all the time, and tend to offset each other. They bring no net increase or decrease of the total money supply. And not only do they leave the volume of the money supply unchanged, but ordinarily they do not alter its make-up. Individual changes tend to offset each other and the proportion of cash to deposits remains the same. The case obviously is the same if a depositor does not himself cash the check, but another person in whose favor it is drawn does so. Here too it is merely a question of the form of the effective money supply.

Second, and more important, is the case where a check is drawn in favor of a person who does not cash it but proceeds to send it

[1] The same outcome of course ensues if the depositor has two banks. If a man has a deposit in bank A and also in bank B, and pays a loan due in bank A by a check on bank B, his deposit in bank B is lessened by so much, but no corresponding deposit arises in bank A. Similarly if he purchases securities from bank A and pays for them by a check on bank B. Again, if a bank loan is found to be bad once for all—if a borrower becomes bankrupt—his deposit in the bank, whatever it may be, is simply cancelled, the debt being treated as disposed of.

to his own bank "for deposit" (which in the eye of the law means for collection). Then the sum appears as a deposit in that bank to the payee's credit—an available means of payment for him, and the check goes through the clearing house back to the drawer's bank. The payee's deposit account becomes larger; that of the drawer becomes smaller by the same amount. The total of deposits has been neither increased nor decreased. But that total of deposits had been in the first instance made larger by the creation of the drawer's deposit. So far as the immediate parties are concerned, the second deposit is not indeed a created deposit; it is a lodged deposit. None the less it is the result of the earlier creation; the deposit persists as an element in the country's medium of exchange.

This kind of outcome is representative of the way in which the creation of deposits leads to a net enlargement of the volume of money. It deserves to be considered somewhat more in detail. Of course the borrower whose account has been credited with a created deposit wishes to use the money. He is likely to draw checks soon, which then lead to lodged deposits by the payees. He is not indeed likely to denude his deposit completely, except perhaps for a very short time. As has just been said, there is always an understanding with his bank, more or less explicit, that he is to keep a "fair" balance; commonly in the United States a balance on the average of something like 20 per cent of the original loan. He may reduce it to that minimum immediately after the loan is made, may even get below it for a while. But as time goes on, his business operations lead to receipts as well as payments. He makes deposits in the bank as receipts reach him, these being obviously lodged (not created) deposits. As the maturity of his loan approaches, he will lodge more and more, his deposit account grows, until finally he pays his loan,—and then the deposit is cancelled. True, it may be cancelled only in name. If he is a good customer, can show to the bank that his business is doing well, and can present a satisfactory balance sheet and income account the loan is readily renewed. Even if his affairs are not in exemplary shape, the bank may still think it wiser to renew —to "nurse" him—rather than to insist on repayment and thereby

perhaps bring on the customer's collapse and a loss for itself.

Business men in good repute get accommodation year after year, deposits being steadily created (and renewed) over and over for them. Some are dropped—the less successful—and their deposits disappear. New customers come in. All have lodged deposits as well as created deposits. It is not easy to make out with any precision what proportion of an existing total began as the one kind, how much as the other. A very large part—probably the greater part—originated by creation, not by lodgment. The two combined constitute a huge mass of purchasing power, "check money," which is steadily available, is steadily being used, flows through the banks and the clearing houses. It is easily subject to increase and decrease, yet remains for periods of considerable length much the same. It enlarges when business grows more active. In times of acute crisis or prolonged depression it may show marked and rapid decline; and then the decline means that debts have been paid, or marked off once for all as uncollectible,—that deposits have been cancelled.

The term "lodged" deposits may be used, in a narrower sense, to indicate those which are the result of a lodgment of actual cash. The largest quantitative lodgment of this kind comes from the retail dealers who day by day send their receipts to the bank for safety as well as for convenience in eventually drawing checks against them. Another lodgment is that of deliberate savings which are put away in cash form, more especially by persons of small incomes. Commercial banks are almost always glad to receive such savings deposits. In England they receive them as "time deposits," and in the United States as "savings deposits." In the United States they more often go not to the commercial banks but to savings banks which do nothing except to act as agents for their safe-keeping and investment. It is better that lodged deposits of this last kind should be in such special institutions and separately safe-guarded by law.

§ 3. It is sometimes said that this process of continued creation and re-creation, of lodgment and circulation, means that the same money is lent by the banks over and over again,—eight or ten times. The phrase is not accurately descriptive of what goes on, and it seems to me ill judged because it implies that there is

nefarious practice. The banker who hears it is puzzled and resentful. What he sees is that people are "bringing their money to him" all the time. Even when he starts with a loan, his conception of the operation probably is that he has first lent "money" to a borrower, and that this person then has left more or less of it with him on deposit. He sees no meaning in his everyday operations other than that people leave their funds in his care, and keep on doing so day in and day out; and he may allege with some pride that this is done simply because they trust him.

What is of course true and is in the minds of those using the phrase which offends the bankers, is that the quantity of deposits created and set going by a bank has a limit in the monetary resources which it holds,—the legal tender money in its till and the reserves equivalent to cash held elsewhere, say in a Central Bank. Commonly enough the till money and the reserves of a bank amount between them to something like 10 to 12 per cent of its demand deposits. Hence it can be said that loans and deposits are built up on this foundation in the ratio of 10 to 1. But it is neither accurate nor rightly significant to say that this same reserve money is used over and over again. The process is not so simple, and its economic meaning is rather obscured than clarified by the phrase. What happens is that not the reserve but the deposits go around over and over again, as all money does. The whole elaborate system grew to unexpected and vast dimensions because of its convenience for depositors and its profit for bankers. It did not arise by evil intent or evil doings, nor on the other hand did it bring unmixed good. The underlying question here, as in most questions of social and economic development, is whether the pursuit of gain within the bounds set by existing law and morals works on the whole for the general good. On this aspect of the matter, the important one, more will be said presently.

§ 4. Essentially similar to the creation of deposits, yet in outward aspect different, is the arrangement common in England for "overdrafts." In the United States an overdraft means that a person who has a deposit account draws checks to a greater amount than stands to his credit. Any check so drawn may not be honored; if presented by a stranger, it is likely to be returned with the nota-

tion "no funds." If both drawer and payee are well known at the bank, and the overdraft is modest in amount, the drawer is notified at once of what is supposed to be an inadvertence and requested to make good. A stigma of ineptitude or even of dishonest intent attaches to an overdraft by an American; and no bank wishes to let it appear in its statements that overdrafts have been allowed to any appreciable extent. But an English bank deliberately arranges with a customer that he may overdraw up to a given amount, say £1000; that is, he can draw checks up to that amount and they will be honored even tho he has no formal deposit. He can deposit at any time thereafter and what he deposits serves to lessen his overdraft. He is charged interest for the net amount of overdraft, for the week or the month.

As soon as a transaction of this kind is initiated, the customer has at his command purchasing power up to the full stipulated amount, and is in precisely the same position as the American for whom a deposit is created. The difference is that the Englishman is not charged with interest until he begins to use the credit put at his disposal. Another difference is that while a deposit is entered at once on the bank's books and is accounted for in its published statements, an overdraft rests on a more or less informal arrangement between the customer and the bank manager. Only when the customer draws a check does an overdraft enter the bank's accounts. The item "overdraft" is a familiar one in British bank statements and causes no lifting of eyebrows as regards either party in the transaction.

So far as concerns the stock of "money" in the community, the two arrangements come to the same thing. In either case that stock is increased; in either the bank's customer has check money which he can use as readily as cash in his pocket. But the fact that the item of overdraft enters into the bank's obligations to an unpublished extent makes it impossible for an economic statistician or a public official in England to know just what is the amount of the circulating medium which exists in this form.

§ 5. All use of cash in bank payments would cease if there were but a single bank, if all deposits were kept at it alone, and if all payments were made by checks on it. The payee of a check ordi-

narily "deposits" it. Should each and every payee keep his account in the supposed single bank, it would deduct so much from the amount credited to him who drew the check and add so much to the amount credited to the payee. No cash would pass, and the payments would be effected simply by substituting one person for another as the bank's creditor (i.e. depositor).

Suppose now that there are two banks, having different sets of customers. Some customers and depositors of bank A will draw checks payable to those of bank B; and on the other hand customers and depositors of bank B will draw checks payable to those of bank A. Each bank will receive daily checks drawn on the other bank, deposited with it for collection. The banks can readily offset claims, and arrange to pay only such differences as may be due to one or the other. It is conceivable that they may even arrange not to pay the differences at all, but to keep a running account, the balance being one day in favor of bank A, the next in favor of bank B, with a tendency to equalization in the end. At all events, the amount of specie (or like cash) that will have to pass between them from time to time will be small as compared with the transactions concluded by mere offsetting.

Next, instead of two banks, imagine a dozen, twenty, any number; the same process can be carried on. Daily each bank receives checks drawn against the other banks. Daily each has to meet the checks drawn by its own customers and deposited by the several payees with the other banks. Tho a few checks may be disposed of within each bank (when the payee happens to keep his account in the same bank as the drawer), most of them are not settled quite so easily. Yet they are disposed of in practically the same way. They go thru the clearing house, where the process of offsetting checks against each other is carried to its farthest limits.

A clearing house is a general organization of the banks of a given place, having for its main purpose the offsetting of cross obligations in the form of checks. Every bank sends to the clearing house all checks it has received against other banks, and each has to meet the checks drawn against it. The totals, of course, exactly offset each other. Any one bank may have a debit or a credit balance, and is in strictness bound to pay in cash the balance due or

entitled to receive it in cash. In practice, the balances are often settled in other ways. Sometimes they are paid by checks on another large bank. In London, clearing-house balances are settled by checks on the Bank of England, with which each of the associated banks keeps a deposit account; and then all the payments are finally effected by mere offsetting, without the use of any cash at all. In those American cities in which there are Federal Reserve banks, checks on the Reserve banks are used; and as between these several institutions an offsetting method is again applied. In the interior parts of the United States, clearing-house balances are often settled in New York exchange; that is, by checks (commonly called "drafts" when drawn by one bank on another) on large banks in New York with one or another of which each interior bank keeps an account. The final settlement of the transactions then takes place through the New York clearing house, with a minimum use of cash. Sometimes clearing-house balances are simply left to go from day to day as debts due by one bank to another, subject to payment of interest by the debtor bank. A bank which is debtor at the clearing house one day may expect to be creditor another day, and unless it happens to have an unusual supply of cash in its own coffers may let any daily balance against it stand as a debt due the creditor banks. This practice of course depends on the willingness of the creditor banks to let the debt stand, and also upon the rules agreed on by the banks in general for clearing-house operations. The healthier practice is to maintain payments of balances once for all. But the same reasons which induce all deposit banks to reduce their cash to the minimum lead them to mitigate by postponement the strict obligation to settle clearing-house balances in cash.

Most of the checks are disposed of at once by being balanced against each other. The larger the banks, and the greater their number, the more likely it is that any one will have at the clearing house about as much to its credit as to its debit. In a comparatively small city it is more likely that the offsetting will not be complete, and that any one bank will have a considerable balance in proportion to the total transactions to receive or pay. In a large city the offsetting process is achieved with extraordinary completeness.

In New York and London 95 per cent or more of the clearing-house exchanges are settled by offset, the balances payable and receivable by individual banks amounting to less than 5 per cent of the total. Practically the same proportion is found in cities like Philadelphia, Boston, Chicago, Liverpool, and Manchester.

Clearing houses develop with the growth of deposit banking. Deposit banking, again, has grown most in English-speaking countries, and most of all in the United States. The London clearing house was established in 1775; this early date is conclusive proof that deposit banking was then carried on in large volume by a considerable number of banks. The New York clearing house was established in 1853; a date somewhat late in view of the early and rapid development of deposit banking in New York. Every considerable city in the United States now has its clearing house, and they exist in not a few cities that are inconsiderable. In 1908 there were 115 cities having clearing houses.

The final outcome, in the countries where deposit banking is fully developed, is that practically the whole of the immense volume of transactions is settled by mere offsetting. "Cash" is needed by banks in their ordinary doings only for the purpose of providing their customers with pocket money. The chief call is for pay rolls, and cash which goes out of the banks in large lumps for this purpose returns to them thru the retail dealers, in the same amounts as it goes out and almost as soon. Five per cent of deposits is quite as much as a bank needs to keep on hand in the form of till money. When the use of banks as repositories of all spare funds, and so the use of deposits as currency, has become highly developed, the rest of the circulating medium becomes involved in the all-pervading operations of the banks.

§ 6. The part which paper money, whether bank or government, then plays in the deposit-using countries becomes substantially the same as that which subsidiary coin has long had. Subsidiary coin always is a minor part of the circulating medium, and is treated as a thing by itself. It never pretends to be full weight; it is never coined freely; it is manufactured and put into circulation in amounts determined by the calls which come to the mints. Any one can get such amounts as he wants in exchange for

other money. The total of subsidiary money in circulation automatically regulates itself in response to the community's occasions for small and moderate payments. And that total—here is the important point—is determined *by* the general price and has virtually no effect on the price level. I say, virtually; meaning thereby that this constituent of the money supply is so small a proportion of the aggregate that its effect on prices is almost negligible. The great item is deposits; next come the legal tender notes of governments and their banks (gold coin has become as negligible as subsidiary coin); last, subsidiary coin.

Notes have gradually come to be, not indeed negligible as compared to deposits, but a factor of much less weight. It is the deposits which are dominant in the relation between the total money supply offered for goods and the aggregate of goods which come on the market to be sold for money. The other money items trail behind. They are like the cavalry in modern warfare; still much to the front, esteemed, much talked about, and of aid to the main body of troops; but their numbers are restricted and they serve chiefly as a supplement to the main body, and count little in the decisive juxtaposition. Ordinary hand-to-hand money conforms in its volume to that of the main body, the deposits, and the volume of these is the dominant factor in the determination of the price level.

Monetary theory and monetary policy have been but slowly and haltingly readjusted to this remarkable development. Monetary theory has had to face a complication that could be neglected in former days. Under the simpler conditions of the past, the supply of money could be treated as an independent variable. It could be so treated when specie was the decisive constituent; the volume of specie being dependent on the product of the mines. Paper money also could be so treated when the stage of firm public control was reached; the policy of government determined how much there should be. Deposits, however, as they came to be created, presented a new kind of problem. They arise in direct connection with the volume of goods which come on the market. For the other kinds of money it can be said—at least for short periods, even periods of moderate length—that their quantity is fixed by causes in which the quantity of commodities plays no part. But for deposits

it would seem that the very appearance of goods leads to the creation of deposits as the goods are dealt in; there may be a prompt, even immediate increase of the money supply *pari passu* with the increase of the goods supply. What limits are there to this creation of deposits? Is it to be controlled? A new question of public policy thus arises. To all these matters attention will be given in the chapters to follow.

§ 7. The command of ready money and so of capital which deposit banking brings about has come to be of growing importance in modern times. It has aroused more and more of critical consideration and has led to proposals for radical control by legislation.

The problem is not so new as is often supposed. Before the middle of the nineteenth century it was already apparent in the United States that ordinary commercial banking, which quickly came to be predominantly deposit banking, was not in fact confined to commercial operations or "liquid assets,"—that is to the short-time borrowings of traders and manufacturers. The discount of their paper was (and commonly still is) expected to be a marshalling process; a steady succession of short-time loans, regular repayments and then new advances. Such the operations of the banks largely were and largely are. It is true that in practice the constant repayments were from the start no more than nominal. Not only were loans often renewed as they became due but it was commonly a part of the understanding between banks and borrowers that they would be. But side by side with these were some loans which from the start had no liquidity, being based not on sales of goods marketed or soon to be marketed but on long-time ventures and commitments. Legislation in the United States repeatedly tried to prevent this widening and deepening; it was desired to keep banks within the narrow channel thought sound and proper. The attempt was unsuccessful even in the days when banking was on a moderate scale. It came to be even less so as time went on and banking came to be on a larger scale and more complex. The tendency in the nineteenth century and still more in the twentieth was toward great banks which combined with "purely" banking business that of the financing and promoting of permanent capital

investments. The tendency toward greater size combined with greater complexity appeared conspicuously in the great German banks of the nineteenth and twentieth centuries. In the United States the part played by long-period commitments of the deposit banks emerged at the very beginning, induced as it was by the eager demand for capital in the development of a new country's resources. This trend toward the ready use of the deposit machinery for investment operations became marked in the early decades of the twentieth century.

To make clear what this means, what complications and difficulties arise, something must again be said on a more general matter: the chain of steps in the process of investment and of capital construction.[1] The individual whom we think of nowadays as an investor commonly has little to do with the making or the management of physical capital. He is merely the purchaser of something which represents a capital equipment already constructed by some one else. Even in the early stages of the modern period he was likely to do something analogous when he lent on a mortgage of real property,—a pledge of urban sites on which houses or workshops already stood, or of agricultural land which had been cleared and smoothed. Nowadays he does it more simply and easily for himself, but less simply as regards the connection with "real" capital. He buys stocks or bonds, securities representing fractional ownership in some physical outfit which exists already, usually as a going concern. The real capital had already been made at the instance and under the leadership of promoters, initiators, semi-speculative adventurers. These are the active agents in the growth and diversification of that large and increasing part of the community's equipment which is managed by others than its owners. Usually the promoters are closely allied with investment houses and with the large private banks which act mainly as middlemen for investment. Each composite group of this kind is again allied more or less closely with a deposit bank, perhaps with more than one. Loans are made and deposits created which at the very start put into the hands of the promoters a large part of

[1] Compare with what was said in Chapter 5, and what is said later in Chapter 38, Vol. II.

the financial resources needed for carrying on the substantive operations.

Savings are thus *anticipated.* They are brought to bear, so to speak, before they have matured. To use a phrase made familiar by Mr. Keynes, investment outruns savings.[1] It cannot do so indefinitely. The promoters, investment houses, deposit bankers look forward to an eventual settled stage when securities can be offered to the ultimate savers, the permanent investors. If things go well, the securities are gradually absorbed. Not all of them necessarily get into other hands; for the promoters or their heirs may retain some portion. But in general they are widely distributed. The promoters and the banks wind things up, and expect to do so with a profit. If things go wrong, a loss must be written off by the promoting group and by those more adventurous investors who have put their savings into the securities at an early stage. The deposit banks themselves may have to face a loss. Often a considerable time elapses before it is clear whether there is success or failure. Often enough, too, even when failure is indicated to be inevitable, the dark prospect is concealed or ignored and an endeavor persists to "get out" by still marketing the securities. A good fraction of them is likely in such cases to remain still in the hands of the promoters and their banking associates, and this in time of crisis may become the occasion for their financial undoing.

Such is the procedure by which in modern times investment is keyed up and hurried up. Not only speed but progress and diversity are stimulated. Doubtless it is some dim awareness of the nature of the operations that has caused the term "promotion" to be applied to them. Industry is turned to new ventures ahead of time: it is promoted. A similar play of language appears also in the use of the term "securities" to designate the symbols of investment which are passed on to savers when a venture is supposed to have reached a settled and secure stage. As a matter of fact the procedure inevitably entails risks, greater or less according to the stage of entry;

[1] I use Mr. Keynes's phrase, but with a different meaning of "investment." He means by it the mere purchase of securities; whereas in the text here it means the making of physical capital.

and the securities may or may not be secure. The uncertainties and risks are even greater than those which in any case arise from the mere lengthening of the period of production. In promotion, the characteristic element is novelty: inventions, technological improvements, new kinds of management and organization, not least the development of newly discovered natural resources. And it is a characteristic accompaniment that success and failure alike are closely connected with the fluctuation and cycles of industry and the recurring booms and depressions. The business cycle is an accompaniment of the rapid advance in material progress which marks the modern world.

§ 8. The changes in the range of deposit banking operations which began in the nineteenth century and in the United States became more pronounced in the course of the twentieth were accompanied by a change in the kind of men at the head of the larger banks.

The management of a purely commercial bank is not on the whole very difficult. It calls for prudence, probity, adherence to routine and system. The banker of the older type must be a good judge of enterprises and a good judge of men. He must be well informed about what is going on in the community about him. He should be the confidant and friend of his customers. He must have caution, and refrain from long-time commitments. Given these traditions, men of good intelligence and character may conduct banks of deposit with success for generations and build up for the bank a reputation and prestige—a sort of trade-mark value—which will lead to growth in size and in profits.

The more complex operations of a financing and promoting bank call for rarer qualities; not only for judgment and discrimination, but for high business shrewdness and for venturesomeness. Risks are taken, heavy commitments made to new enterprises, plans laid for a comparatively distant future; not least, the right men selected for the conduct of new enterprises. This was the nature of the operations of the great private banks of older days; such houses still play a large part in the origination and promotion of investment. The large part which banks of deposit have come to take in these operations has led to a change in the men at

their head. They direct and speed up business quite as much as merely aid it in running smoothly along the accustomed channels. The risks are greater; and since a larger constituency than before is connected with each bank—a multitude of non-borrowing customers as well as a great range of active new concerns and enterprises—the penalties for error and failure are spread over a wider circle.

CHAPTER 24

DEPOSITS FURTHER CONSIDERED. RESERVES, RUNS, REGULATION

§ 1. Cash in the vaults of a bank tends to be reduced to the minimum. But its resources must be largely "liquid"; commercial paper, salable securities demand loans.—§ 2. Relation between the rates of discount on bank loans and the reserves held by the banks. Demand loans and speculative transactions.—§ 3. Runs and panics.—§ 4. Far-reaching effects of panics.—§ 5. Repeated breakdowns in the American banking system. Peculiarities of monetary and banking conditions.—§ 6. Further remarks on the peculiarities of deposit money.—§ 7. Regulation for security of deposits in the United States. The great step, insurance, taken in 1935.

§ 1. AGAINST their demand liabilities in the form of deposits, banks must have cash or assets readily convertible into cash. Bank notes, while they have remained demand liabilities in form and in law, have been so completely regulated or taken over by governments that they no longer give concern to the issuing banks. But deposits are always regarded by the creditors—the depositors—as being in fact payable on demand in cash, i.e. in legal tender money, without notice or warning. It is for these that the bank must have a supply of cash on hand or obtainable at once.

Cash in a bank's vaults is "idle" money; it is earning nothing. The bank is under constant temptation to reduce to the lowest terms its holdings of specie or other legal tender money. Some cash it must hold, in order to meet demands at the counter and balances at the clearing house (unless, indeed, the latter can be met in other ways). Additional cash it may hold against the contingency of a "run"—sudden demands by depositors resulting from a distrust in the bank. But as a rule this possibility is little regarded unless under compulsion by the law. Why not invest "idle" cash, put it out in loans or purchases of securities, and get an income? Hence the actual holding of legal tender money tends to be reduced to the minimum which experience shows to be needed in the ordinary course of business. That minimum is sur-

prisingly slender. Five per cent of the demand obligations seems to be ample. The English banks of deposit, which issue no notes, and do not trouble themselves (for reasons to be explained in the next chapter) about any reserve against runs, rarely keep more than this proportion against their deposits, and often do not keep so much. American banks also, unless compelled by law to keep more (as they commonly are), find that for their day-to-day operations they get along comfortably with five per cent.

None the less, the other resources of a bank must be of a sort to enable it to face the fact of great demand obligations. It must have quick assets. Its loans are on short time, and in a well-managed bank are marshalled in such a way that some of them are maturing week by week and day by day. Recurrently, and at short intervals, the bank becomes entitled to repayment of loans, and thus expects to be in a position quickly to increase its cash or diminish its demand liabilities (i.e. its deposits).

The typical form of the short-time loan, as has already been said, is discounted commercial paper. All manufacturers and all wholesale dealers, and most retail dealers, give credit to purchasers and in turn borrow from the banks. Loans to them, on short time and in connection with their current dealings, are to a high degree secure; for to meet them is the first condition of a man's standing in the mercantile community and so of the very possibility of maintaining himself in business. By the older tradition, the banker was the confidential friend and adviser of the business men who were his customers; well informed upon their affairs, and disposed to aid them with credit according to their pecuniary deserts. This sort of relation, with discount of commercial paper based on it, still plays a large part in ordinary banking operations.

Side by side with these intimate relations there have always been transactions of a more cold-blooded sort, and of late these have grown in volume and importance. Loans are made on collateral; that is, on the pledge of property that can be sold by the bank if the loan is not promptly paid. Securities—stocks and bonds of all sorts—are the most welcome form of collateral, because the stock exchanges give the most perfect facilities for disposing of them. Every bank has a certain proportion of loans, commonly secured

by stock exchange collateral and payable on demand, which it is expected at once to convert into cash if there should be any sudden large presentation of demands against the bank by depositors.

Other assets of a quickly realizable sort are also found in a bank's resources. Familiar and easily salable securities are often held, such as can be turned into cash at a moment's notice in case of need. Every English bank parades on the published balance sheets its consols, and treats these as if they were the same thing as cash. The United States bonds and the state and municipal bonds which our banks enumerate in their published statements are of the same sort. When a bank, after it holds such securities to the amount consistent with general safety and repute, still finds that it has more cash than is needed for its current operations, it will "go into the market" and buy with the surplus anything that seems safe and profitable. It may buy ordinary "good" securities, even tho in its usual course of business it does not buy and sell stocks and bonds. It may buy "outside paper"; that is, the promissory notes of business firms which are not its own customers and depositors. This is done thru note brokers, a set of middlemen who have become of growing importance in American banking. The note brokers peddle the commercial paper of well-known firms among banks whose resources are temporarily, sometimes permanently, greater than they can utilize among their own clientele. The practice of buying such notes of course increases the range of the banker's cold-blooded operations. It is said to distribute the banker's risks better; he is not so much bound up with the fortunes of the particular clique or trade that makes up his regular customers. On the other hand, it makes him deal with persons whose affairs he knows with little certainty; and it brings possibilities of overexpansion by the borrowers and of unexpected loss to the bankers.[1]

§ 2. It is the business of banks to lend. They lend more freely in proportion as their cash holdings or reserves are large. Hence

[1] Purchases and sales by ordinary commercial banks of commercial paper and of securities on the "open market" are not to be confused with "open-market operations" of the Federal Reserve banks. The latter represent an instrument of monetary control and are of peculiar significance because conducted by the central banking system. See Chapter 27, § 6.

there is a close connection between the rate at which bank loans are made, that is, the rate of discount, and the quantity of money held by the banks.

A common notion is that the rate of interest depends on the total quantity of money in circulation, falling when that is plentiful, rising when it is scarce. The notion, thus broadly stated, is unfounded. The abundance of money affects its exchange value—the general range of prices. Advocates of paper money have often supposed that an increase of the quantity of money would lower the rate of interest. In fact, the period during which prices are rising commonly is one not of lower interest, but of higher. When once the definitive stage of higher prices has been reached, there is nothing in the nature of the situation to affect the rate of interest.

But tho the rate of interest does not depend on the quantity of money circulating in the community at large, the bank rate for loans is affected by the quantity of cash which is held in the bank's vaults. The commercial world is constantly speaking of the value of money and the plentifulness of money; and it uses both phrases in a special sense, referring to the banking situation. By value of money it means primarily the rate of interest or discount on short-time business loans. By plentifulness of money it means a relative abundance of cash in the banks, leading to a free offering of loans. Relative abundance, be it noted: a large or small supply relatively to the demand obligations of the banks. When they have more cash than seems needed for daily calls and for safety or prudence, they lend freely. Thereby they either add to their demand obligations (by increase of deposits) or part with some of their cash in the direct purchase of commercial paper or securities. In either case the relation of cash to liabilities is altered, until something like the normal situation, or supposed normal situation, is reached. Conversely when their cash is scant as compared with the demands reasonably to be expected, banks draw in. They refuse to make new loans, or to renew old ones; tho they will "take care" of their steady customers, they turn a deaf ear to others. Hence the rate of discount varies according to the plentifulness of cash held by the banks. Large cash holdings lead to "easy money" and free lending, small holdings to "tight money" and guarded lending.

These tendencies and the fluctuations in interest rates which result appear most conspicuously in the case of demand loans. A demand loan, it must be remembered, is subject to call by either party: the debtor can be called on to repay at any time, but he also has the option of making repayment at any time. When banks have plenty of cash, they lend freely, and at very low rates, on demand; for if other more profitable ways of using their funds should present themselves, they can at once call for payment of the demand loans and turn to those more profitable. Hence one hears of 'call" money in New York, where these fluctuations are most striking, quoted as low as one per cent a year. On the other hand, a trader who is in stress for means to meet immediate liabilities will pay a high rate for a demand loan, knowing that he can repay at any time and hoping to do so in a few days. At the height of a crisis, call money has been quoted in New York at 100, even at 200 per cent a year. Of course no one would borrow at this ruinous rate thru a year; but it might be done for a few days in order to tide over an emergency.

Demand loans are usually cold-blooded, made to any one on the deposit of collateral. The debtor must pay when called on, and if he fails to do so, the collateral which he has given is sold at once.[1] The loans are commonly connected with transactions on the exchanges, especially on the stock exchange, and are an essential part of the mechanism by which speculation is facilitated. From the point of view of the bank, they are a highly convenient part of its business. A profit is certain, sometimes large, sometimes small, but always appreciable; and yet the bank's resources are not tied up and cash can be called back—at least by the individual bank—if there is more profitable use for it elsewhere, or if there is a sign of danger. And from the point of view of public advan-

[1] Many loans which are nominally on demand are in reality not subject to this sort of drastic treatment. Loans in this form are made more than in previous times to merchants, taking the place of sixty-day or ninety-day paper yet not essentially different from it. It is understood that the bank will not in fact "sell out" the borrower. And even stock exchange call loans, when made to brokers who are regular customers, are payable on demand more in name than in fact. Banks like to parade on their published accounts a large volume of demand loans, as if their promptly realizable resources were abundant. In fact, the convertibility into cash is often more nominal than real.

tage, there is also some gain. For sundry useful business transactions funds are wanted over short but uncertain periods, and for them demand loans are adapted.

But there are also grave public disadvantages from such transactions. They facilitate gambling speculation, not only in stocks, but in grain, cotton, and other standardized commodities. All the real and serious evils of speculation are made greater, or at least made more easily possible, by the willingness of banks to lend funds on call to any one who will pledge sufficient security. With the tendency toward specialization in all modern industry, there are some banks which turn more freely than others to this sort of lending; and indeed in every great financial center there are a few banks which make this almost their exclusive business. Such lending, too, is closely connected with the tendency to accumulate all spare bank cash in the financial centers, like New York and London; a tendency which is more particularly connected with the development of deposit banking and with some of the dangers in that system, of which more will be said elsewhere.

The rate of interest on ordinary commercial loans—time money for thirty days, sixty days, ninety days—of course shows much less fluctuation than that on demand loans. For regular customers and depositors of banks, the rate of discount often varies very little, even tho cash in the banks may be more or less abundant; there being an understanding that they shall have "reasonable" accommodation at a "fair" rate, that is, at the customary or current rate. The rate on advances of this sort goes up and down somewhat, and oscillates about the general rate of return on permanent investments. Discounts for less regular customers fluctuate more sharply according as the cash holdings of the banks are larger or smaller. In times of stress such loans may be very hard to get and may be made at high rates—8, 10, 12 per cent; while regular customers may still be taken care of at the traditional rate, say 6 per cent or 5 per cent. Conversely, when money is "easy," the banks may buy "outside paper" on terms which yield them less than the usual rate. The business man, in arranging his banking connections, often has a choice between these two ways of securing his credit. He may deal steadily with one bank, very likely a conservative

one, paying to it a fairly steady rate of interest thru good times and bad, and sure of support in periods of stress. Or he may float his paper thru note brokers, and borrow here and there at varying rates. Then he takes his chances as to support in the periods when no bank has much free money and when many business men are demanding loans. The former course is that which conduces to the safe and steady conduct of industry: the latter is that which promotes the recurrence of commercial crises. Yet the latter often seems immediately the more profitable and, under skillful as well as venturesome management, may be in fact the more profitable. In every community there will be found two types of banks and the two types of business men. The fluctuations in the rate of interest naturally appear most sharply in the dealings between the venturesome banks and the venturesome business men.

§ 3. In general, then, deposit banking implies that the banks have a great volume of demand liabilities and a comparatively small amount of cash with which to meet them. If there is a general and sustained run on all the banks, the cash almost inevitably proves insufficient. There is then nothing left except a general suspension of cash payments. To prevent such a general run, to maintain the confidence of depositors, to keep in working order this intricate part of the machinery of exchange—these are among the main objects at which legislation and the policy of banks have been directed in the English-speaking countries.

When any one bank is beset by a run—caused perhaps by some unfounded rumor, some unreasoning fright among its depositors —it appeals for aid to the other banks. These have the strongest motive for granting aid by supplying cash from their own holdings; since fright is contagious, and the failure of any one bank is likely to precipitate a general run. But the condition on which aid is granted usually is, and always ought to be, that the bank in straits be solvent; that its loans and other assets prove on examination to be sound, and sufficient in the ordinary course of events to meet its liabilities. The possibility of a run, and the necessity in that case of exposing its whole situation to critical professional eyes, are the strongest forces for preventing reckless and dishonest banking. A bank which is once fairly going may be really insolvent

and yet able to keep going for a considerable time as if all were well. It can carry on its books, as if good, loans or securities which are bad. So long as depositors continue their daily round of deposits, loans, checks, there is little to reveal the true situation. But once there is a run, the bank must show its hand. Where there is an organized clearing house, a committee representing this institution (that is, the combined banks of the place) examines the threatened member, and learns whether aid is deserved. If it is, the reserves of all the banks are massed at the point of danger. Every depositor in the imperiled institution is told he can have his cash if he wishes it; and at the same time public assurance is given by the clearing-house committee that the bank is solvent. But if it is not solvent, and must be wound up with possible loss to depositors, the combined banks face the situation boldly, "take care" of the embarrassed depositors, and endeavor to quiet general apprehension. By such means an incipient panic may be averted.

It must be said that in the banking experience of the United States, this sort of emergency has not always been courageously faced. A bank that is known to be insolvent has been kept going because the others are afraid of the aftermath from its open failure. If it is large and conspicuous, as may well be the case, there is fear that its failure will precipitate a panic and runs on all the banks. Moreover, the question whether that bank, or others more or less in similar plight, is really insolvent, is not answered at once by an inspection of its books and possessions. There will be sundry loans and sundry securities which may be good and may not be. Only with the lapse of time will it become clear what the situation really is, whether failure is inevitable or recovery can be eventually attained. The rivalries of banks among each other, personal friendships and antagonisms, a lurking fear of some among the banking managers that their own institution would not come well out of a searching examination,—these may be important factors. Wisely or unwisely, it may be decided that a sinking bank, even one that deserves to be scuttled, shall be saved in the fear of a general collapse.

If a general panic and a general run do develop—if moreover, some banks are really insolvent and others are in an uncertain con-

dition—the situation is more difficult to handle. Here it is unquestionably a vast advantage if there be some one great strong institution with ample cash holdings and unshakable prestige. For the banking institutions of the United Kingdom, the Bank of England has become in such times the citadel of refuge. It can undertake to supply cash when needed, and to guarantee solvency if there be real solvency. A similar position is held by the Federal Reserve banks of the United States; tho at this writing (1939) the procedure to be followed in times of stress has not been worked out and settled as fully as in England or its effectiveness in allaying panics as well proved by experience. To prophesy that acute financial distress will never recur in England would be unsafe, and it would be at least equally unsafe to make such a prophecy for the United States. But the unreasoning, mob-like panic has become highly improbable. On all this more is said in the later chapters on Central Banks.

§ 4. The financial panic is often, tho not invariably, the first stage in a general overturn of the business situation. On the later stages and the much more difficult questions of the business cycle, more will be said in another connection.[1] I confine myself here to the banking and monetary aspects.

All business men conduct their affairs on the basis of giving and taking credit. Each individual is both creditor and debtor, has his bills payable and his bills receivable. In the ordinary course of things, these obligations are met punctiliously. Failure to meet them means that the delinquent loses his standing in the business world; he is no longer in the game. It is on this ground of expediency—of honesty being the best policy—that the discount of mercantile paper is so secure a banking investment.

Anything which unsettles the expectation that mercantile debts will be promptly met may cause a panic among business men. Each knows that his paper is coming due and that to enable him to meet it he must receive payment of what is coming due to himself. If he fails to pay his own obligations, he gets poor comfort from the fact that his own failure is due to the failure of his debtors to pay; his standing is broken none the less. Now all obliga-

[1] See Chapter 13, Vol. II.

tions are likely to become greater and more dependent each on the other during an upward industrial movement. Where there has been some really serious maladjustment, failures are inevitable. But then it becomes also possible that one failure will entail another, and this still another, until business firms topple over in succession like a row of bricks. Of this sort of collapse a dramatic example occurred in its simple form during the great crisis of 1857; both in the United States and in England an extraordinary number of firms then collapsed, chiefly because of all-pervading panic.

When the storm is brewing, the one thing needed in the business community is assurance against indiscriminate ruin. This can be given by the banks, if they are themselves in a position to render aid. What merchants and manufacturers want at such times is "accommodation." They do not want cash. It is true, as will presently appear, that there may be at the same time a run on the banks for cash. But while some business men may join in the run, it rarely touches the mercantile community at large. What is needed for its peace of mind is primarily the assurance that support will be afforded against possible temporary embarrassment. Loans are wanted, not cash; or rather, assurance that loans can be had if needed. Business men want assurance of being "taken care of." In deposit-using countries, they want the banks to make them advances—to credit them with deposits—which can be used in meeting their own accruing obligations even tho the debts due to themselves fail to be met promptly.

§ 5. In the United States there was nothing in the experience of the nineteenth century and the first decade of the twentieth to indicate that the financial panic was a thing of the past. The situation was in many ways different from that in England and in many ways contained greater elements of danger.

During the second half of the nineteenth century, banks of the large cities, and especially those of New York, occupied, it is true, a position analogous to that of the Bank of England. But the analogy did not reach far. They were many in number, and tho combined for some purposes in Clearing House Associations they could not act with the energy and promptness of a single institu-

tion. Even if they had been organized to act unhesitatingly, they were not in a position to give all the assurance and support that were needed. Their reserves of cash were only such as the national banking laws compelled; sometimes a little more, but at the times when panics were likely, seldom appreciably more. Not least, they were themselves not above suspicion. It is true that as a rule most banks are solvent and even supersolvent. But there are commonly some black sheep, with rumors and suspicions of more. Banks in general, solvent or insolvent, are uneasily conscious that they are not invulnerable; their position necessarily ceases to be so when public confidence begins to be shaken

To these causes of danger was added the fact that deposit banking in the United States was extraordinarily widespread. And it remained so. Not only was the total volume of deposits in the United States very great, but the number of individual banks and of individual depositors was and is enormous. There is a larger proportion than in England of persons who are likely to be affected by unreasoning panic. Deposit accounts are kept not only by those doing business on a considerable scale and by persons of large means, but by small tradesmen, farmers, women. These easily get into a fright when some great bank fails and rumors are flying thick about others. An overt run or a silent steady withdrawal of cash may then be precipitated. The banks on the other hand are scattered, are sensitive to the possibility of sudden demands, and are themselves by no means free from panicky feeling. Many of them are small; many, large and small, conduct their operations in ordinary times with a bare minimum of cash. When danger threatens, those outside the great cities telegraph urgently for cash to the city bank in which they keep a deposit. They do so not only to meet real drains by their own depositors, but to provide against possible or anticipated drains. Among the banks, as among their individual depositors, a spirit of *sauve qui peut* may develop, and then a full-fledged panic may burst.

These conditions brought about repeated breakdowns of the American banking system. On four conspicuous occasions, in 1873, in 1893, in 1907, and in 1933, complete collapse ensued. In each of these great panics the banks of the country virtually suspended

payments. Thereby they committed acts of bankruptcy, and under the strict letter of the law could have been forced into liquidation. The fact that suspension was universal and well-nigh inevitable caused its strict legal consequences to be ignored; and after a few weeks or months the usual course of payments was resumed. But during these weeks and months, on all four occasions, legal obligations were put aside. Neither individual depositors nor depositing banks could get cash which they had the legal right to demand. No doubt their demands were in one sense unreasonable. Individuals called for cash because they wished to hoard by tucking it away in drawers or in safe deposit boxes. Outside banks wanted it partly because their own depositors made similar demands, partly because they themselves were in a fright lest such demands should come. Whatever the cause, the breakdown was well-nigh complete. A depositor on these occasions might be allowed to draw pocket money—a few dollars—from his bank; but any demand for considerable sums was met, in most cities and by most banks, with flat refusal.

The panic—acute stage of a crisis—does not last long. A few weeks of excitement and anxiety, of banking and mercantile collapses, of pressing demand for "money" (i.e., loans) at high rates of discount, are followed by rapid subsidence and quiescence. Almost invariably, cash accumulates in bankers' vaults within a few months of a panic and the rate of discount falls to a low figure. These conditions hold for a considerable period, longer or shorter according as the revival of activity comes late or early. During this period the banks, tho willing and able to extend advances, find the business community unresponsive, and an abundance of cash in their hands goes hand in hand with low and falling prices.

§ 6. Deposits in their operation as a circulating medium are among the most marvellous of economic phenomena. Like the division of labor which they serve to facilitate, they developed by no intention and for generations had little restraint or guidance from legislation. They work out their results by processes but half understood by the very persons who manage them, the bankers. In countries where deposit banking is largely resorted to, like England and the United States, all wholesale transactions, and a large

and growing volume of retail transactions, are carried on by them. Deposits combine in a remarkable degree safety and convenience. They are less subject to theft than cash because a check is payable to a specified person, and the bank is answerable for making payment to that person only or to his endorsee. They are convenient because a few strokes of the pen serve to remit any sum, however large, and to remit that sum precisely to the last cent. The mechanism is wonderfully smooth and effective in its working.

Two things are essential for the development of deposit banking; or rather two phases of one thing,—confidence. Checks cannot pass from one person to another unless there is confidence in the probity of the drawer. Business custom has supplied substantial ground for this confidence. Quite apart from criminal penalty, he who draws a check without having credits at the bank to meet it commits industrial suicide. More important is confidence in the bank itself. The basis of the entire system is the repute and good standing of the banks. It can endure, for any individual bank, or for the banks as a whole, only so long as people think the bank's obligation to pay cash to be equally good as the cash itself.

Given these conditions, the vast but frail mechanism maintains itself in unceasing round: loans are made, deposits created, checks drawn, deposits renewed, loans again made, and so on. It consists of a mere mass of debts, contracted without any formality, and evidenced only by a few figures on bank ledgers and pass books. It is a sort of phantom circulating medium, ever appearing and disappearing, never substantial, always in danger of melting away from a breath of suspicion, yet so serviceable as to be renewed after every collapse and to be maintained notwithstanding every danger.

§ 7. The repeated unhappy experiences of the panics of 1873, 1893, and 1907 led finally in the United States to the remaking of the currency and banking system in 1913. A later collapse of essentially the same kind in 1933 brought a further step of far-reaching effect in 1935. I shall reserve for later consideration the main features of the Federal Reserve system as established in 1913

and modified by later legislation. In the present connection it is enough to note one part of the supplementary legislation of 1935 —that by which a system of insurance was created with the object of guaranteeing the safety of deposits. At the end of well nigh a century of experience, the conclusion was reached not only that bank notes but bank deposits also were part of the circulating medium; and that depositors no less than note holders were entitled to protection from loss by bank failure.

That holders of bank notes were entitled to protection against loss in case of bank failure, had been settled in the United States by the establishment of the national bank system.[1] That system also sought to give protection to depositors by requirements for reserves, bank examinations and reports; but the protection proved illusory. The fundamental question was rarely discussed: is it desirable that special provision be made for the protection of the holders of bank notes? Or should note holders and depositors be put in the same position?

No doubt the difference in treatment rested historically on the fact that the similarity between notes and deposits was not perceived. Tho deposits subject to check formed part of the circulating medium quite as much as notes, and indeed were quantitatively much more important in countries like Great Britain, the United States and Canada, they were not commonly regarded as "money"; yet notes were so regarded. The usage of everyday speech, reflecting like all such usage an established general tradition, goes far to explain the special protection long provided for note holders.

Some real grounds for special security to note holders are these. In the first place, notes are more likely to be held by the poorer and dependent classes. Deposits are used chiefly by the well-to-do. Notes circulate among all classes, and notes of the smaller denominations are likely to be in the hands of workmen and others of slender means. Next, and not less important, is the difference in the way in which a person becomes a creditor of the bank. A depositor almost always becomes creditor by his own choice; a note holder commonly becomes so without any volition of his

[1] See Chapter 22, p. 303.

own, and moreover by a process of whose legal import he usually knows nothing. A note circulates from hand to hand as "money." The person to whom it is offered in payment would commonly find difficulty in refusing it. Ordinarily he is quite unaware that in taking it thus freely, in the eye of the law he is simply replacing another as creditor of the issuing bank.[1] Which individuals are the note-holding creditors at any given time is a matter of accident; since each person receiving a note keeps it until he has occasion to use it in a purchase. A depositor, on the other hand, selects his bank with some deliberation. No doubt, he is often influenced by the bank's mere propinquity or by its general reputation. None the less, the initiative comes from him and the first responsibility is his.

These distinctions, however, must not be pressed too far nor be permitted to obscure the fundamental point of resemblance—that deposits, like notes, constitute part of the *de facto* circulating medium. The same fundamental reasons which make it important that notes should be secure, make it important that deposits also should be secure. The essential question concerns the expedient ways of promoting security.

In the legislation of the United States, the principle of providing in some way for the protection of depositors was early established. It showed itself not only in those requirements as to reserves of national banks which have already been noticed, but in a whole code of banking legislation. The nature of the loans which a national bank might make was rigidly defined. Not only

[1] The legal position of the payee of a check is different from that of the holder of a bank note. The payee of the check does not, like the note receiver, become at once a creditor of the bank. The bank's liability is only to the drawer of the check (the depositor). If the bank refuses to pay the check when presented, the depositor only, not the payee, has a right of action against it. On the other hand, the sending of a check in payment of a debt does not at once liquidate the debt. Should the bank fail, or for any reason refuse to pay the check, the debtor who had sent it is still liable. If, indeed, the payee of a check fails to take steps with reasonable diligence for its presentation at the bank on which drawn, the legal situation becomes different. If he puts the check away, and waits unduly before presenting it, he substitutes himself for the drawer as creditor of the bank. Failure of the bank in the interval then means loss to him, not to the bank's original creditor (depositor). Hence the business practice of always sending all checks received for immediate "deposit," i.e. for collection at the clearing house. By this process the receiver of a check makes himself as promptly as possible a creditor of his own bank.

in the national bank law, but in those of the states, there was and still is regulation of the extent of loans to any one individual and of loans to directors; not least, provision for publicity of accounts and periodical examination. The Comptroller of the Currency at Washington has a staff of examiners and exercises large powers over the national banks; the several states either have similar bureaus for their own banks, or are moving toward such supervision.

All this regulation was unique in the United States. Just as the requirement as to cash reserves for deposits was unknown elsewhere, so was all the detailed regulation of loans, reports, and special liabilities of officers and directors. The situation was a curious one. In a country where the general tradition had been to let capitalistic industry pursue its course unfettered, the very center of capitalistic operations was subjected to a degree of control undreamed of in other countries. The cause of this remarkable extension of state interference is to be found partly in the early development and wonderful spread of deposit banking, but still more in an underlying dim consciousness that here is really a most important and far-reaching part of the circulating medium. Once the system is fully established, no individual of the property-owning and property-managing class can keep out of it. It is indispensable that he have his bank of deposit and his bank account. And tho he may choose his own bank and may be supposed to be on the watch as to its character and solvency, his means of getting information are necessarily uncertain. The public concern in banking, which at first was chiefly for the security of notes, came to be no less for their equally pervasive and far more powerful successors, the deposits. Hence the proposal that deposits should be made absolutely safe, like notes, was not an illogical or revolutionary one. Obviously, no method of segregation of particular assets (such as may be used in regard to notes) can suffice for the purpose; since the only possible security for all deposits would be the solidity of all assets. The only feasible method is some sort of insurance—compulsory contribution by every bank to a public (or publicly supervised) guarantee organization, out of which the deposits of a collapsed bank would be met. The main

objection to this proposal is that one great safeguard—perhaps the greatest safeguard—against reckless banking would be removed, namely the banker's fear of the depositor. If every depositor knew that his "money" was sure to be forthcoming in any case, being guaranteed by the state or other adequate organizations, pressure on a bank from uneasy depositors would be less likely to follow suspicious doings. The probability of the presentation of demands on suspected banks would be greatly diminished. A bank might pursue a reckless course for an indefinite time, or at least for a longer time than if the confidence of the depositor needed to be constantly nurtured.

This objection, tho strong, is not necessarily conclusive. Reckless banking takes place now, even under the eye of the unguaranteed depositor. If the guarantee were one not of immediate payment but only of ultimate payment—if the depositor, tho secured from eventual loss, were still subject to the possible inconvenience of having his funds "tied up" for a time in a liquidating bank—it would still be to his interest to be watchful, and to withdraw his account when suspicious.

In conclusion, I venture to quote what I wrote on this topic in 1921:

> The course of legislation on this matter, as on others, is likely to be much affected by actual experience. A succession of conspicuous bank failures, bringing great loss to depositors, would immensely strengthen the movement for deposit guarantee. Much depends, too, on the development of the relations between the Federal Reserve system and the scattered banks which remain outside that system. And not a little depends on the general trend of social and political development. Banks and banking constitute the most characteristic feature of the régime of private property and private enterprise; and as that régime comes to be modified in greater or less degree, there will be likelihood of further and more elaborate regulation of banking operations at large.[1]

The expected has happened. The great banking collapse of 1933 led to a system of deposit insurance. In 1935, an important feature —in my opinion the most important—in the act revising the

[1] *Principles of Economics*, 3d ed., Vol. I, p. 387.

Federal Reserve system, was the establishment of a federal organization (the Federal Deposit Insurance Corporation, owned by the government) which provides insurance against deposits for all banks within the Reserve system and for a large range of other banks as well. Deposits were thereby made secure in the sense of a guarantee against ultimate loss, tho not against immediate inconvenience. It remains to be seen (1939) what effect this momentous step will have on the banking structure as a whole. It may easily lead to the elimination of the excessive number of petty banks, and eventually to a wider use of branch banking. But the first and fundamental result will stand: bank depositors will be protected against the calamitous losses of older days.

CHAPTER 25

DEPOSIT BANKING (*Concluded*)

§ 1. The way in which the general welfare is promoted by commercial banking operations: continuity of industry, selection of the more capable among business men.—§ 2. Investment operations of commercial banks. Deposits lodged by persons not active in business.—§ 3. Creation of deposits and the march of material progress. Criticisms of the deposit structure and proposals for its radical change.

§ 1. WE proceed now to another aspect of the relation between the various forms of banking operations and the general welfare. Wherein do banks promote prosperity? Often enough they are spoken of as mere leeches, parasites living on the body public, contributing nothing to its well-being. And conversely their doings are often discussed as if the only question were *how* they are carried on and how they are related to the general monetary situation, with little regard to their bearing on the good or bad working of the whole economic mechanism. On this range of subjects, something has already been said;[1] but its importance is so great and there is so much misunderstanding that a further consideration is permissible, even tho it involves some repetition.

Commercial loans, whether made to a bank's regular customers or to others thru the purchase of outside paper, have economic effects that are easily followed and are readily seen to be good. They make the community capital more effective by making its savings go farther. The doings of an individual merchant are likely to be irregular. Those of manufacturers are also uneven; more uneven if they "make and sell," less so if they "sell and make." Both are affected by seasonal changes of demand, and by ups and downs which it is difficult to foresee. If every dealer and producer had to rely on funds or savings of his own, he would have too much available at most times, but barely enough and often not at all enough at peak times. Bank loans smooth the irregularity for the individual concerns, while the banks themselves can maintain

[1] See Chapter 23, §§ 7, 8.

comparative regularity because the varying demands of different concerns and industries tend to offset each other. Banks thus promote the continuity of the use of capital. So far it is all to the good.

At least equally important, but not so obvious, is the influence which banks exercise on the make-up of the business world. As will appear in later chapters, something like a process of natural selection determines who shall come to the lead in the conduct of business. In that selective process the banks play an important part. They lend freely to the man of whom they think well; they turn a deaf ear to him whose projects seem unpromising. Their willingness to give credit makes it possible for the capable man to extend his operations, even tho he start with little or no means of his own. No doubt they sometimes make mistakes, and enable empty pretenders or reckless speculators to get command of large resources. But bankers, to repeat, need above all things to be good judges of men. On the whole they put the control of the industrial forces into the hands of those likely to turn them into profitable channels. This holds good more especially of commercial banks of the older and simpler type. The heads of such banks are usually well informed about the men and doings of their town or city. Not only do they commonly make their loans for short periods, but they keep in touch with what is going on and are able to judge when it is well to make or renew loans, when to refuse. Judgment of men counts more than anything else,—which beginners to nurse, which ambitious schemers to handle warily. This sort of judgment is of cardinal importance in all banking operations; not only in the simpler ones of the kind here described but in those more complicated which will be presently considered. Of commercial banks it can be said with the minimum of qualification that their characteristic operations—short-time loans to merchants and manufacturers—tend to put and keep the control of industry into the hands of those most likely to conduct it well.

But even a bank which is as purely commercial as a bank can be does not confine itself to such operations as these. It is likely to have deposits from persons not in active business. These are obviously lodged deposits. Often they are merely checking accounts maintained for convenience in paying bills. As the de-

posits come in, and as spending of incomes by their holders takes place with a lag after their receipt—sometimes with a long lag—the bank finds itself with more cash on hand than is needed for even the most conservative conduct of its business. It buys securities or outside paper, or lends on collateral to others than its regular customers. Substantially the same is the case of a bank which has large deposits from business concerns, drawn on as convenience or necessity may require but little subject to sudden withdrawal. If such lodged deposits are large and steady, as is often the case with a long-established and well-reputed bank, it will buy and hold long-term bonds of a kind which can be marketed readily and turned into cash. In so doing, it acts substantially as a savings bank; savings of a quasi-permanent character are turned at once into the channels of investment.

§ 2. We turn now to that other kind of connection between deposit banking and investment,—the creation of deposits—which develops incidentally from the more familiar practices. The phenomena appeared at an early stage in banking history, but by the middle of the nineteenth century gradually came to be of a size and range not dreamed of before.

First, a word of reminder. It is often said that a bank supplies capital and by so doing adds to the productive resources of a community. In the sense of physical capital a bank obviously does nothing of the kind. Tools, machinery, materials, are created primarily by labor, seconded by the aid from saving and lending. Yet a bank, tho it does not create physical capital, is a most important instrument in regulating the *command* of capital and in promoting the effective use of capital. As regards savings banks, investment concerns, and similar financial institutions, nothing needs to be added to what has already been said;[1] it is simple and obvious that they are intermediaries in the process by which saving leads to real investment and the construction of capital.

Commercial banks have often been described as performing the same functions in the same way. They are said to gather rills of savings—money means not immediately needed by the owners and deposited by them in the bank—and to lend them to producers;

[1] See Chapter 5.

very much as savings banks lend the accumulations which are more deliberately set aside. So far as deposits are lodged, arising by the actual putting into banks of spare funds, this description fits. But so far as deposits are created by the banks, it has none. Here command of capital is supplied by the banks without that sort of saving which is ordinarily associated with the process of investment. Money means are created, and the command of capital is supplied, without apparent cost or sacrifice on the part of any saver.

§ 3. As has been explained in Chapter 23, the creation of deposits means a hastening of economic growth. The deposit bank system has served greatly to promote the rapid industrial development of the English-speaking countries and especially of the United States. The extraordinary growth of this country during the second half of the nineteenth century and the earlier decades of the twentieth was made quicker and more sweeping. This growth was more than welcome to the public at large; it was a source of pride and congratulation. It became also the goal of the countries of the Continent, even tho they did not fall in with the monetary system which was so signal a feature in the United States and England. The seamy sides of it all were not perceived, or if perceived were accepted as concomitants of the grand march of progress.

In the period that came after the war of 1914–18, and its aftermath in the depression of the fourth decade, the ideal of growth piled on growth lost its glamor. So it was with the deposit structure as part of it all. Especially in the United States the question was asked whether the pace had not been too fast and too dangerous, entailing consequences for the distribution of wealth and income and for the even march of industry which must be offset against the gains. With this general attitude arose the question whether the creation of deposits and all that hangs thereby had not better be put an end to once for all. The most radical suggestion for change was that commercial banks should be compelled to keep reserves up to the full amount of their deposits,—a reserve of 100 per cent.

I cannot undertake here an examination of this scheme in all

its aspects. To carry the plan thru would involve a complete overturn and remaking of banking structure and tradition. History shows clearly, I think, that the final outcome of legislation on banking is difficult of prediction, if not impossible. This is the more the case when the changes are of a revolutionary kind. What the American economic system would be like if this far-reaching change were put thru, who can say? Above all, would the pace of material progress be lessened? Or if not lessened, is it certain that it would take a better course, or at least a smoother one?

A fundamental question underlying all such discussions is that of the adequacy of the present volume of production. Is there enough turned out per head or per family to make it quite acceptable that we should go on as we are and not think it important that the national dividend be increased? The present organization (disorganization?) is vastly more effective than that of a century ago; is there no occasion for making it still more effective? Surely there is. We can increase welfare by more abundant production and we should arrange our economic structure with an eye to doing so. But the forward movement involves a process of trial and error. It means not only irregularity but failures as well as successes. The question is whether the improvement of that situation step by step, in such way as to lessen irregularity yet maintain progress, is better than to substitute once for all something radically new and untried.

So it would seem to be with all the fundamental questions of monetary and banking policy; they are part of the problem of social and economic reform at large. There is no panacea in sight. The best procedure is by tentative steps which—so far as we can see—are in the right direction. Such is the nature of the legislative and administrative steps taken of late (the period 1933–39) for the reshaping of the Federal Reserve system of the United States. And such seems also the nature of the steps being taken thruout on the reshaping of the entire economic structure.[1]

[1] Compare Chapter 27, in this volume, on the Federal Reserve System.

CHAPTER 26

CENTRAL BANKS; FRANCE, ENGLAND

§ 1. Three great systems to be examined, those of France, England, and the United States. In all three countries, the Central Banks, while private corporations in form, are in effect agencies of the government.—§ 2. The Bank of France, has a complete monopoly of note issue. Importance of this, in view of the comparatively slight development of deposit banking. —§ 3. Wide permeation of the Bank's operations. Rediscounting and the effects. The Bank essentially a banker's bank.—§ 4. Close connections of the Bank with the public finances. In times of stress, an agent for putting out inconvertible paper money.—§ 5. The Bank of England. Its organization. Absence of branches. Complete separation of notes and deposits. Issue and Banking Department. The Issue Department and the fiduciary issue. Bank notes and currency notes.—§ 6. The Banking Department. Relation to the other banks. The Bank's reserve. Bill brokers.—§ 7. Working of the system in times of crisis, use of the discount rate, open-market operations.—§ 8. The Bank and the government. Enlarged note issue and inflation in 1914–18. New functions after the war, and especially since 1931. Growing concern for industrial as well as monetary stability.

§ 1. CENTRAL Banks, originally planned to be primarily agents and aids for governments, have gradually been called on to perform wider and more important functions. They have come to be the sole issuers of bank notes, and these notes have come to be virtually government paper. While in their early stages they were not essentially different from other banks—merely the first among equals—they have come to be unique in position: bankers' banks, the leaders and supporters of the others, even their rulers. And this has inaugurated a further stage, at which they are beginning to be called on to assume new and larger functions: the maintenance of stability in the price level and, not less, of stability in trade and industry. It is an old story that the coinage and printing of money has a pervasive influence on the whole of economic life; an old story too that banking and credit profoundly affect its course. That Central Banks as arms of the government undertake to exercise some deliberate control over it all is new.

It will serve best to bring about the main questions and prob-

lems that arise if we examine in succession the history, traditions, working of three great systems—the Bank of France, the Bank of England, and the Federal Reserve banking system of the United States. Different tho they are, they are confronted by similar problems and are shouldered with similar responsibilities.

In one important respect these banks are alike: in form they are private corporations. They are not owned and managed by the governments as are post offices and the like. The power which the government exercises over central bank operations is not direct, but indirect. The indirect relation—the domination by government of a private corporation—is largely a matter of tradition, being the continuation of a type which arose when the banks in the main were mere fiscal agents. Whether there is wisdom in retaining this sort of mechanism has been much debated. A common opinion, of much weight, is that the corporate form of organization makes it easier to keep a central bank aloof from momentary political exigencies and to maintain a steady settled policy. It is not without significance that of late years, in the course of the widening range of government operations, this plan—a corporation controlled by the state yet not part of the political structure—has been resorted to in fields where no such traditions had developed as in the field of central banking. The experience all around has done something to indicate that when a government enters a new industrial sphere in a community where most industry has been and is private, the familiar tested forms and methods of private business are best suited for large continuous operations.

§ 2. The Bank of France is the simplest in structure and in range of operations. As a private corporation it has its stockholders, directors, appointed officers. Public control dominates because the important executive officers, the Governor and two sub-governors, are appointed by the government. These manage the Bank. How far they are independent of the Board of Directors is not easy to say. That Board, fifteen in number, was long elected in a peculiar way: not by the entire body of stockholders but by those 200 of them who held the largest number of shares of stock. This curious oligarchic arrangement, maintained without change for over a century after the establishment of the Bank by the first

Napoleon, brought it about that the Board had great continuity of personnel and character. On financial and business matters it probably conduced to sagacious and cautious management.[1] There were well-established traditions and a general conservative temper which the public authorities could not disregard,—the more so in view of the short tenures of French ministries.

§ 3. The Bank of France has a monopoly of note issue. That privilege is of far-reaching importance, for several reasons. First, there is no other paper money; the Bank's notes occupy the field. Next, these are virtually government paper; usually convertible into specie, but on several occasions made inconvertible. The government authorized, nay compelled, the suspension of specie payments when war came in 1870–71, and the notes remained inconvertible till 1878 (tho not depreciated from 1873 on). It again made them inconvertible when war broke out in 1914; and they remained so till the devaluation and resumption of 1926 (*de facto* 1926, *de jure* 1928). About this step, a landmark in monetary history, more will be said in other connections.[2] At this point it is enough to note that the Bank's notes, whether or not convertible into specie, have all the qualities of government notes. They are legal tender for debts, circulate without a sign of cavil or hesitation, and bid fair to do so indefinitely.

Finally, and what is most important, the Bank's notes constitute by far the largest part of the circulating medium. The practice of deposit banking and of payment by check has taken little root in France. It has established itself only in Paris and a few other large centers, and in these only for a limited circle of bankers and wholesale dealers. Most transactions, large as well as small, are settled in Bank of France notes. Hence deposits form only a minor part—usually less than a quarter—of the demand liabilities of the Bank. Against them, as against its notes, the Bank is required to hold a minimum gold reserve of 35 per cent.

A substantial part of the deposits at the Bank of France consists of balances held for the other French banks. These other banks

[1] The change made in 1936, by which the directors were to be chosen by a "democratic" franchise, was not likely to make the Board less conservative, but lessened its disposition to influence the government's action and its power to do so.

[2] See Chapter 28, § 11.

generally find it convenient to keep a part of their reserve resources in this form and they constitute the larger part of the Bank's deposits. One consideration is that such deposits enable them to use checks drawn on the Bank of France for purposes of clearing. Another is that they are enabled to make use of the widespread facilities of the Bank for the collection of obligations due them. The main consideration is that deposits with the Bank afford the most effective means of access to further supplies of notes, whenever extra cash is needed either to meet calls arising in the ordinary course of business or to meet unusual calls.

The operations of the Bank of France in one sense permeate the entire country more than those of other central banks do, and yet in an important respect less so. Its notes circulate thru the length and breadth of the land, and its branches and agencies are widespread. In addition to the central office and several branches in Paris, it has branches—altogether over a hundred and fifty—in all the principal cities and towns. The branches, tho under close supervision of the head office, are permitted to perform most of the ordinary operations of banking. They discount bills, make loans, accept deposits. The Bank maintains also nearly five hundred minor agencies and offices in all parts of the country. These do not accept deposits; their main purpose is to facilitate the making and collection of payments which arise in the course of the financing of business, not least among which are those arising in connection with the principal and interest of the public debt.

Nevertheless the Bank's connection with the business community at large is not so close as all this would indicate. The limitation is the result—very likely a fortunate one—of an old statutory requirement, strictly enforced, to the effect that commercial paper in order to be discountable at the Bank must have three "names" —i.e. three persons liable for its payment when due. The ordinary bill of exchange—and this form of commercial paper remains the usual one on the Continent for domestic as well as for international paper—has by its nature only two "names," those of the drawer and the acceptor. The Bank cannot discount a bill unless there is a third name (indorser). This is commonly the name of a private banker, who first discounts to the borrower and then re-

discounts at the Bank. Often he does not rediscount at an early stage in the operation but waits—according as he himself may need the command of ready money—until the time approaches when the bill comes due. The Bank rediscounts an enormous number of such bills having but a short time to run. Not only are the pieces many, but on the average each one is for a small amount. The Bank thus finances at its branches myriads of credit transactions for persons of all kinds and yet comes into direct contact with few of them. In this regard its relations with the public are different from those of the great private banks of England, Scotland and Canada. These also have a large number of widely scattered branches. Not being fettered by law as the Bank of France is, and being able to extend loans directly thru the creation of deposits, their connection with the business community is quick and direct.

Hence the Bank of France, notwithstanding the permeation of its facilities, remains essentially a banker's bank. It is the court of last resort for all the banking institutions of the country. The clearing house payments of the large cities, and above all those of Paris, are made thru checks on it. All large banks keep deposits with it, and rely on it for aid in times of pressure or distress. The habitual resort to it for rediscounting gives it a power of control over the banking system as a whole. Its rate affects the entire borrowing market, by its direct dealings with some business concerns (chiefly larger ones) and with the great private banks, and also by the less direct procedure of rediscount. When its rates are low, rates thruout are low; when high, all rates are high. On the whole, the tradition of bank operations in France is that the discount rate shall be low, and shall change slowly; a tradition which the Bank is able to maintain because it is in the position to create that form of credit money—notes—which dominates the circulating medium.

§ 4. As with all Central Banks, the operations of the Bank of France are closely connected with the public finances. This is in one way a matter of no great economic significance, being merely a matter of the handling of the government receipts and payments and of the public debt. But it is of the first importance

in that the volume of the circulating medium is settled chiefly by the Bank's notes. In England and the United States the Central Bank's control of the circulating medium is indirect, chiefly thru action on the deposit structure; and the results of this situation, as we have seen and shall see still more, are not of a simple or predictable kind. In France the quantity of notes is the main thing. They were, indeed, long limited in total amount by the law; but the amount permitted was enlarged not only with the growth of population and wealth but also—what proved of enormous importance—when a war brought sudden fiscal needs. In two great periods of stress, 1870–71 and 1914–18, the Bank was authorized to put out a much greater volume of notes than before, placing the money at the disposal of the government. The notes were made inconvertible; the Bank in effect became the agent of the government for issuing inconvertible paper. In both periods the process was less sudden and erratic, and probably the volume was less than would have been the case if the paper had been put out by the government directly. In 1870–71 the Bank was able to be of service to the country in this way without sensible derangement of the price level. So skillfully did it manage the operations on that occasion that no noticeable depreciation set in, and specie payments on the old basis were resumed without difficulty in 1878. Very different in degree, tho not in kind, were the results that ensued during and after the war of 1914–18. The government again borrowed from the Bank on a vast scale; in response the note issue was multiplied many fold. Prices soared to four or five times the pre-war level. For a decade the French monetary supply consisted in the main of a huge mass of inconvertible paper, with a maximum limited by law but always in danger of enlargement by new legislation. After the war the heavy government borrowings from the Bank gradually ceased. Ten years after its close (1928) notes were again made by law redeemable in gold, but with reduced content of the gold franc. Probably the inflation had been kept within bounds that were comparatively moderate (in view of the prolonged and enormous strain) by the circumstance that the notes were those of the Bank, not government paper. But it was still so great that, as has already been noted,

the franc was finally devalued and a much higher price level in terms of "francs" was accepted once for all.[1]

§ 5. The Bank of England is the oldest and most famous of the great modern banks. It resembles the Bank of France in the important particular that it has long been the foundation of the country's entire monetary structure and the source to which both public and private finance turn in times of stress. It differs from the Bank of France and from all other central banks in its organization, its traditions, its ways of working. Most important of all, the monetary system which it supports is different.

The Bank of England, to repeat, is a private corporation. Its shareholders, like those of any other private undertaking, receive dividends. The shareholders, moreover, elect the board of directors —known as the "Court"—and the directors in turn choose from their number the executive officers. These consist of the governor, the deputy governor, and the comptroller. The governor and the deputy are appointed to office for a term of only two years; the comptroller holds office for life, his responsibility being by tradition that of assuring a continuity of policy. Thus the Bank of England is managed, so far as the letter of the law goes, quite without government intervention. But the established customs and traditions are no less binding than the law, and are in this case of economic consequence at least as great as if they were established and defined by law. By long experience the Bank has come to appreciate fully its public responsibility, and now conducts its affairs in essentially the spirit of a public institution.

In structure the Bank of England is different from any other central bank. It operates no branches (in the usual sense) and has no other office than that in London. Its organization is regulated by the Bank Act of 1844 and by later legislation which follows

[1] The maximum of the bank note issue was 1,800 millions of francs at the beginning of the war of 1870–71, and had been raised to 3,200 millions when specie payments were resumed in 1878. The corresponding figures for the war of 1914–18 were 12,000 millions in 1914 and 41,000 millions in 1928. There is no longer a limit to the total note issue of the Bank of France, the former legal provision having been repealed by the Monetary Law of 1928. The same law provides that the Bank shall hold a gold reserve of 35 per cent of the combined amount of its notes in circulation and its liabilities on current account.

On the problems of devaluation, compare with what is said below, Chapter 28, § 11.

the principles of this,—*the* "Bank Act." The salient feature is the complete separation of the issuance of notes and the management of deposits. Notes are put out by the Issue Department, which has that function only. Deposits are managed by the Banking Department, and in this the real business, or at least the main business, is conducted. To all intents and purposes the two departments are separate institutions.

The operations of the Issue Department are strictly limited. Bank of England notes are legal tender money, and their issue is carefully regulated by law. Up to a specified amount notes might be put out under the Act of 1844 without any gold cover. Against these notes the assets held by the Bank—which has free choice in the matter—consisted mainly of government securities. For every note beyond this amount the Bank had to hold pound for pound in gold. The specified uncovered amount was fixed in 1844 at £14,000,000. It was provided that, as other banks then having the privilege of issue should come to lose that privilege, the limit of the uncovered notes of the Bank should be correspondingly enlarged. The expectation was that the other banks would gradually cease putting out notes, and that the Bank of England would in time secure a complete monopoly of issue. Under this provision the Bank's "fiduciary issue," as it is called, had risen to £19,750,000 in 1923, by which time the old banks of issue had practically disappeared and the Bank had achieved a monopoly of issue.

The fiduciary issue was later increased immensely; another among the far-reaching changes brought about in so many directions by the war of 1914–18. Until 1914, no Bank of England note could be issued for less than £5—a restriction which, since it caused gold coin to be resorted to for a great number of transactions, limited very much the use of notes. During the war, however, the government stopped gold payments by the Bank, thus abandoning the gold standard. Gold coins were withdrawn from circulation and government paper issues of denomination less than £5 (so-called currency notes) were put out by the Treasury to take their place. When the gold standard was restored again in 1925 a change of vital importance was made; it was decided to give up the use of gold coins altogether. Accordingly, under the Cur-

rency and Bank Notes Act of 1928, which amalgamated the Treasury note issue with that of the Bank, the restriction on the denominations of bank notes was removed and the Bank was authorized to print £1 and 10s notes. Notes were made redeemable in gold, but only in bullion and when presented in large blocks. Gold could still be had for industrial use and (more important) for shipment abroad. But gold was no longer coined. The new notes supplanted the former sovereigns and half-sovereigns. To require the Bank to keep 100 per cent in gold against this added liability would have imposed a serious and quite unnecessary burden. The limit of the uncovered issue, therefore, was raised to £260,000,000. At this figure it stood in 1936.[1]

The principle upon which this kind of regulation of the note issue originally was founded can be stated in brief terms. There is a certain minimum volume of notes which always will find a ready circulation, at any given stage of the price level, because they are indispensable for hand-to-hand payments. It is this amount which is held in mind when fixing the fiduciary issue. So far, notes can safely be issued without gold cover because they will never be presented for specie. The total volume of notes in circulation had always been much beyond the fixed limit of uncovered issue. But the notes issued over and above this limit had been backed in full by gold and remained so. As under the Act of 1844, they remained under that of 1925 in the nature of certificates of deposit. Conceivably these notes might on occasion be presented for specie. Since, however, the full amount of specie is held for their redemption, convertibility is assured.

A system of this kind is entirely appropriate for a community where paper notes hold a predominant place in the money supply. But the situation is not so simple when notes are no longer the dominant element. In a deposit-using community such as Great Britain has been so long, regulation of the note issue does not suffice for monetary stability. To understand how far and in what ways the British system conduces to this end, it is necessary to

[1] After January 15, 1938, the limit was to be £200,000,000.

Something more will be said presently (in Chapter 30 below) on the practice of substituting paper entirely for gold in hand-to-hand payments and on the importance of this practice in the working of the gold standard.

examine the position and operations of the more important part of the Bank of England, the Banking Department.

§6. The Banking Department is the center of a great system of deposit banking. England and Scotland are permeated by banks of deposit. The number of these, formerly considerable, has been much reduced by a series of consolidations; a tendency toward banking units of large size which is observable in other countries also, and was accentuated in Great Britain during the war years 1914–18. The total volume of business and of deposits has not declined but rather risen. The banks utilize to the full the machinery of checks and clearing. All of them, many or few, large or small, sail close to the wind as far as reserves are concerned. Here again there is remarkable freedom from legislative control, banks being governed by tradition and custom rather than by law. Part of their resources, often a considerable amount, is invested in consols, which are readily salable; and they have a good deal of "money on call," that is, demand loans. Their actual cash, however, is usually at the minimum of "till money" needed for current demands—often not more than five per cent of deposits.[1] But they do keep, in addition, a certain amount on deposit at the Bank of England and this they regard as perfectly equivalent to cash on hand. Their deposits in the Bank serve to facilitate clearing, checks on the Bank being used for the settlement of clearing balances; they serve also as a resource to be used in case of general uneasiness or of unusual demands upon the banks for cash. These deposits, therefore, act as reserves. They are to be counted, along with what the banks hold in actual cash—which cash is largely Bank of England notes—as reserves against deposits.

There are, then, in the Banking Department of the Bank of England great deposits payable to other banking institutions. These deposits are of more than ordinary importance. Since they are used as reserves by the other banks and are liable to be drawn upon heavily in case of emergency, they need to be amply secure.

[1] Until 1914 that cash was (over and above subsidiary coin) partly gold sovereign pieces, in large part Bank of England notes. After the war, all of it was paper. Gold was thereafter held by the banks only in the form of bullion, and this not for cash purposes but for possible accommodation of customers (bullion dealers) who might want it for transactions connected with international payments, or for use in the arts.

So also must the other deposits of the Bank be secure, those due to its non-banking customers—usually firms conducting large-scale financial operations—and those due to the government. The demand liabilities of the Bank of England must be safe. Yet the Bank is under no legal obligation to keep any specified holding of cash against its deposits. Only by tradition and custom is it bound to keep a reserve or, rather, *the* reserve—the store of cash on which the business community looks as the stay and prop of the entire banking system. That cash is ordinarily kept at between forty and fifty per cent of the demand (deposit) liabilities—vastly more than is usually necessary. The Bank is not managed in this regard primarily for profit. It is managed as a public institution. It keeps a reserve large enough to satisfy all but the most extraordinary demand for cash, one large enough to maintain the solidity of the English monetary system.

The mechanism by means of which this reserve is made available to the other banks is unusual. The English commercial banks, unlike those in most countries, do not as a rule borrow directly from the central bank. Their access to the resources of the Bank of England is indirect. A good deal of their own assets, as was noted above, is in the form of demand loans, made chiefly to the "bill brokers." These firms, often of high repute as regards both means and competence, are especially important in the London money market. They act as middlemen in the handling of commercial paper; usually each of them specializes in that of a given trade or region. They buy (discount) whatever bills are offered for sale on the market and either sell them again to the various banks or hold them until maturity. Their operations are carried on largely by means of funds borrowed on demand from the banks. The bill brokers have the privilege of taking their holdings of commercial paper to the Bank of England for discount, and it is thru them that the banks are able to draw on the cash resources in the Banking Department of the Bank of England. When the banks find themselves running short of cash they call their loans to the bill brokers. The latter, in turn, in order to obtain the cash to pay off their borrowings, take commercial paper to the central institution for discount. The greater the need of the banks

for cash, the larger is the volume of discounting at the Bank of England. It is by this process that the reserve of the Bank is put at the disposal of the entire banking community. When a strain has passed and there is no longer a heavy call for cash, loans to the bill brokers are resumed by the commercial banks and the brokers need no longer present paper at the Bank for discount. As the discounted paper held by the Bank comes to maturity and the collections upon it are made, the reserve is gradually built up again.

§ 7. The operations of the Banking Department are not confined to discounting. In fact, as far as the law is concerned, they are not restricted at all. When the Bank Act was passed in 1844, the Banking Department was considered unimportant. Regulation of note issue alone was deemed necessary. Experience soon showed that this was far from sufficient. It appeared that whenever a crisis occurred pressure was directed on the Banking Department. To this the commercial banks looked for cash; to this the uneasy and hard-pressed financial community looked for accommodation. This had happened before, and is indeed the inevitable outcome of any system of widespread deposit banking. Yet no account was taken, either in the legislation of 1844 or in that of any later date, until the reorganization of 1926, of needs in time of crisis. So capably did the Bank conduct its affairs, so steadily did it recognize in successive crises its public responsibilities, that legislation was not called for. As regards banking operations proper, the Bank still has a free hand to carry out whatever measures it deems advisable. In practice this proved to be an immense advantage, for it enabled the Bank to develop thru experience a strong technique of central banking; a technique by no means perfect, yet probably more nearly so than could have been devised by legislation.

In the rush for safety in times of crisis, banks want cash wherewith to allay the panic-stricken depositors. Merchants want not so much cash as accommodation. They want to be "taken care of." Being in fear of inability to meet existing or accruing liabilities, they want to be sure of loans or extensions of loans; all of which the banks can do for them if panic is allayed. When a

storm is brewing the reserve of the Bank of England is put to full use; discounts increase and cash flows out in the manner described above. Ordinarily this has been sufficient to dispel the prevailing uneasiness. But on some occasions the drain has been so great as to completely exhaust, or almost exhaust, the reserve of the Bank. In the crisis of 1847, and again in 1857 and 1866, exactly this happened. To meet the dilemma resort has been had to the odd but effective device of "suspending" the Bank Act of 1844. The Banking Department, it must be remembered, never has held more than a negligible amount of specie; the bulk of its cash is in the notes put out by the Issue Department. If it runs out of notes, it has under the Act no way of procuring more. But when the Act is suspended and the limit to the uncovered note issue thus removed, the Banking Department can take to the Issue Department government securities and get notes in exchange. The suspension of the Act thus puts additional supplies of cash, potentially unlimited, at the disposal of the Banking Department. This curious device proved to serve the purpose, and so came to be accepted as part of the system. The drain on reserve never required additional uncovered issue to any considerable extent; the mere knowledge that the resource was available sufficed to quell panic.[1]

To maintain its reserve the Bank adjusts its rate of discount, raising the rate when the reserve is undesirably small, lowering it when the reserve is needlessly large. A higher rate tends to discourage discounting and so to curb the drain of cash from the Bank; a lower rate makes discounting easier, and permits cash to flow out freely. More than this, the level of the "Bank rate" affects the entire monetary situation. Partly because the Bank of England holds the basic monetary reserve and can be very gingerly in extending accommodations, partly also because the Bank has attained great prestige, its discount rate becomes a guide for all the banks and financial houses. They adjust their interest rates on loans of all sorts according to the Bank rate. Any movement in

[1] Under the Act of 1926 this relic of the nineteenth century was done away with, the Bank being permitted to issue additional uncovered notes under authorization of the Treasury. This whole series of legislative regulation became obsolete when Great Britain dropped the entire gold standard in 1931. See below, Chapter 30, on the Gold Standard, § 5.

its discount rate affects the pace of lending and the volume of all bank deposits.

There is a second instrument of control: the purchase and sale of government securities by the Bank. These "open-market operations," as they are called, are used sometimes in conjunction with the Bank rate, sometimes in place of it. A purchase of securities by the Bank puts into the money market cash which soon finds its way into the commercial banks. It increases the resources at their disposal and so makes for easier credit conditions. A sale of securities on the other hand tends to tighten credit; it draws cash from the banks and hence makes for higher rates of interest and a reduced volume of lending. Open-market operations are most often used not so much to bring about a positive change desired by the Bank as to counteract other influences which act upon the banking situation; they are used rather to maintain things as they are than to bring new conditions. By combining them with changes in the discount rate, the Bank is able to influence money and credit in ways quite different from the simple limitation of note issue which was thought in 1844 to be the one thing needed. Not that the Bank has power to regulate things exactly as it wishes, nor that it can create and maintain a perfect monetary system; but it is able to impart to the top-heavy and vulnerable deposit structure a high degree of strength and stability.

§ 8. The relations between the Bank of England and the government are in the eye of the law purely of a voluntary sort. The Bank is a private concern, and has dealings with the government as such. It acts as a banker for the government, holds on deposit the funds of the Treasury, and serves as agent for the Treasury in many financial matters. Ever since its founding in 1694, the Bank has been a lender to the government. In ordinary times the loans are on current account (say in anticipation of tax revenue) like those of a bank to any customer small or large. But in times of stress the loans have been of vast amount and have affected markedly the whole financial, monetary, industrial situation. During the Napoleonic wars, the loans were so large and the consequent drain on the Bank's resources so great that the Bank had to suspend specie payments. During the war of 1914–18 specie payments

were again suspended, and the Bank's advances to the government, its note issue, and its public and private deposits reached enormous proportions, all leading to a great inflation. As in the case of the Bank of France, the inflation and derangement were probably less than if there had been direct government paper.

With the opening of the fourth decade of the twentieth century a new and even more important position began to be conceived for the Bank of England and indeed for Central Banks in general. Underlying the new conception was the belief that more might be done and should be done to prevent the upset of the business community by crises,—to aid the government and the public in peace as well as in war. A central bank might become the instrument for maintaining the stability not merely of current financial operations but also that of the price level; and, further—much more difficult and far-reaching—to maintain the stability of the entire industrial structure.

A policy looking to monetary stability, while not entirely novel, naturally gained more and more attention when the gold standard was given up in 1931. Then the question had to be faced on what principles the new money was to be managed. What end was to be aimed at? Stable prices or higher prices? Stable money incomes or changing money incomes? and what kind of banking system was adapted to the end chosen? A policy of the second kind, looking to industrial stability, was prompted by the long and severe crisis that followed the collapse of 1929. In every direction and in every country the call was for public measures to allay this depression and to prevent recurrences; and one obvious way was to improve the monetary and banking structure. New possibilities and new aims for Central Banks were discussed as never before, especially in the English-speaking countries. On these matters something has already been said, and there will be more as we proceed. It is enough to note here that the traditional organization and ways of the Bank of England came to be looked at more critically; with what eventual outcome no one can foretell.

CHAPTER 27

CENTRAL BANKS. THE FEDERAL RESERVE SYSTEM

§ 1. The old National Bank System. The Federal Reserve System; the Board of Governors and its powers.—§ 2. Note issue by the Reserve Banks only; Deposits in the Reserve Banks. Reserve requirements for the Reserve Banks.—§ 3. Working of the System in times of panic.—§ 4. Reserve requirements for the Member Banks. Flexibility of the system in its more recent phases.—§ 5. Rediscounting.—§ 6. Other instruments of control in the Board's hands: changes in the rate of discount; open-market operations; power to alter reserve requirements of Member Banks.—§ 7. Character of the system as a whole, and its future.—§ 8. Central Banks in general. Sole issuers of paper money. New and growing functions and responsibilities.

§ 1. We turn now to the United States. Thru the greater part of its history this country furnished the most important example of decentralized banking. Even under the National Bank System, inaugurated in 1863, there was no pooling of reserves, no true centralization of note issue, no bank for bankers, no unified monetary control. The law laid down requirements for reserves against deposits in the national banks and it stated conditions governing their issue of notes. But each of the national banks kept a reserve in its own vaults, and each issued its own notes. Notwithstanding serious defects in these directions, the system was much better than what had preceded. Notes of State banks disappeared, being taxed out of existence, and those of national banks were assured of goodness and acceptability thruout the country. Standards of banking practice were raised by restrictions imposed in the law and by the creation of a supervisory official, the Comptroller of the Currency. As regards deposits, State banks were not touched at all by the federal legislation and were little regulated by the States themselves. Significant was the legal regulation of the reserves to be held by the national banks; this step squarely faced the fact that deposits are money (check money) and need control for essentially the same reasons as bank notes.

Just how and why the system proved inadequate need not here be considered in detail. It is enough to say that in one crucial respect it proved a failure. It could not stand up under stress. The plan of legally required reserves against deposits did not bring the expected results. What happened was that in times of activity the banks sailed so close to the wind in the matter of reserves that on the occurrence of a crisis they were rarely in a position to meet sudden great demands. Repeated failure in this regard led finally to the establishment in 1913 of a system of central banking—the Federal Reserve System.

The Federal Reserve System differs from the European models of central banking in many ways; most conspicuously in that it is much less highly centralized. Instead of one bank, there are twelve, located in as many principal cities and named according to their location—the Federal Reserve Bank of New York, of Chicago, of San Francisco, and so on.[1] Each of the banks serves its geographical district and each has its own community of customers—"member banks"—among the commercial banks of its district. This arrangement, dictated largely by a strong desire to avoid concentration of power in the hands of a single institution, has distinct advantages: ready availability of the central banking facilities in every part of the country, and ease of adaptation to the differing needs of different regions. The maintenance of a considerable number of branches in other than the designated cities—twenty-five in 1938—promotes the same purpose.

The arrangements as regards ownership and control are unique. The Federal Reserve Banks are not publicly owned, neither are they in the usual sense privately owned. No individuals may be stockholders, and the member banks, who are the stockholders, are *compelled* to hold stock in specified amounts as a condition of membership in the system. The member banks include all the former national banks and a large number of State banks and trust companies, which are permitted (under suitable conditions) to join the system. But the member banks, tho thus the owners of the Reserve Banks, have by no means the same control

[1] The twelve cities are Boston, New York, Philadelphia, Cleveland, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas, San Francisco.

as stockholders in a corporation usually have. True, they elect six of the nine directors of each Reserve Bank; but this is about as far as their powers go. They have no positive authority in determining the operations of the Reserve Banks which are their property, least of all any authority in governing the affairs of the Federal Reserve System as a whole; and their dividends on the stock which stands in their name are limited to 6 per cent, any excess above this going to the National Treasury.

The real guiding and managing agency for the System as a whole, and for its constituent parts also, is a central body, the Board of Governors of the Federal Reserve System,[1] commonly spoken of as the Board. It is composed of seven persons appointed for long terms by the President, one of them selected by him to be Chairman. The Board has been given wide powers over the Federal Reserve Banks, both as regards their executive personnel and as regards their operations. Three of the nine directors of each Reserve Bank are appointed by the Board, and one of these three is designated by it as Chairman of the Directors and Federal Reserve Agent. Moreover, the President and the First Vice President of each Reserve Bank, who act as its chief executive officers, tho chosen by the directors, must be approved by the Board.

As regards operations, the Board's control, somewhat limited in the earlier legislation, gradually came to be great. From the beginning it included the authority to examine the accounts and affairs of the Reserve Banks, require them to rediscount for one another, and "exercise general supervision." It included also the power to review the rates of discount fixed by the several Reserve Banks. The main responsibility in the determination of policy had been left originally in the hands of the several Reserve Banks, because they were thought to be in more immediate contact with the economic affairs of the several parts of the country and hence in a better position to devise and execute appropriate policies. This arrangement was altered by amendments, especially by those contained in the Banking Act of 1935. The powers of the Board by 1935 embraced substantial control over the rates of

[1] This phrase was adopted in the Banking Act of 1935 to replace the former one of "Federal Reserve Board."

discount charged by the twelve Reserve Banks, open-market operations and the other instruments of credit policy. Authority and responsibility were in effect centralized in the Board. It is the position and powers of the Board that give the system the essential features of a great central bank.

§ 2. Under the National Bank System, notes issued by the national banks had formed a large part of the currency for hand-to-hand circulation. Under the Federal Reserve System, a new paper currency, Federal Reserve notes, made its appearance. National bank notes were retired from circulation. Federal Reserve notes became the largest constituent in the supply of cash and the most important. They are issued by the Reserve Banks only after the deposit of collateral security with the Board. The law requires that against the notes a reserve of at least 40 per cent be held by the Reserve Banks and that for the remainder of the notes security be kept in the form of commercial paper or other suitable assets. The required gold reserve of 40 per cent is held at present (1939) in gold certificates.[1]

The Federal Reserve System does not have a monopoly of the entire paper money supply. The United States Treasury issues constitute a large proportion, partly in the form of the United States notes, direct government paper—survivors of the greenbacks of the Civil War; more largely in the anomalous form of silver certificates, in effect a sort of subsidiary currency. The Federal Reserve issues, however, account for the bulk of the paper money supply; moreover, they constitute the one flexible element in the note circulation.

The deposits in the Federal Reserve Banks consist almost entirely (apart from United States Treasury deposits) of those of the member banks and they constitute the reserves of these banks. The member banks are required (1939) to keep against their own demand deposits a reserve amounting to 22¾ per cent in the case of "central reserve city banks," 17½ per cent for "reserve city

[1] The gold reserve, originally held by the Federal Reserve Banks in the form of coin or bullion, is now (1939) held in the form of certificates. By the Gold Reserve Act of 1934 the gold holdings of the Federal Reserve Banks were taken over by the Treasury and gold certificates given in exchange. What will prove to be the final outcome from these tentative moves is not to be foreseen at this writing (1939). See infra Chapter 30, on the Gold Standard.

banks," and 12 per cent for "country banks," and must keep them in the form of deposits with the Reserve Banks. Five per cent is required against time deposits for all classes of banks. This figure had been three per cent in 1913. The deposits of the Reserve Banks include some other items, mainly funds held for the United States Treasury, but also deposits due to foreign central banks, and certain other deposits. The great item is that of deposits which are member bank reserves.

Since deposit liabilities of the Reserve Banks constitute the reserves of the active member banks, these deposits are the foundation of the entire banking structure, and must by tradition be amply backed by the Reserve Banks themselves. The law prescribes accordingly a minimum reserve by these against their deposits of 40 per cent in the form of gold certificates or other lawful money (1939); a minimum large enough to give—so it is designed—great strength to the whole system.

§ 3. The principles underlying the requirement of a large reserve of cash for the Reserve Banks warrants a moment's examination. The system as established in 1913 recognized that the important thing is that not each and every member bank shall keep separately a large reserve, but that there shall be somewhere (in the Reserve Banks) a consolidated strong reserve which can be made available at whatever point it is needed. By this means each and every member bank is expected to have the support of immensely greater cash resources than it could possibly provide for itself. So much of the policy is not open to question. What must also be considered is the precise way in which the required reserve shall be made available to the member banks. If the Reserve Banks are rigidly bound to keep 40 per cent, then under the law they can pay out only what cash they may have *in excess* of this amount. The 40 per cent, far from being always available, then tends to become unavailable by the very process of its requirement.

By subsequent legislation, however, the rigidity which the Federal Reserve System in this regard inherited from the old National Bank System has been so tempered as to avert any serious consequences in times of crisis. The essential principle underlying this part of the legislation was from the start that the Reserve Banks

should be always in a position to furnish ample supplies of cash in times of need without depletion of the required reserves. It is hardly to be doubted that they will always be in that position. And if a crisis should come when there is no excess and the minimum required by law must be trenched on, elasticity is still provided by the power granted the Board of temporarily suspending the requirement, subject to a graduated tax on reserve deficiencies.

It is to be said in general on this matter of crises, panics, and runs on the banks that it is hard to foresee in what way any scheme of legislation will work in practice. When the Federal Reserve System was originally proposed, the one thing on which there was unanimous agreement was that acute monetary crises like those of 1873, 1893, 1907 should be prevented; and the general belief was that at least this much would be achieved by the new System. Yet in 1933 there came a collapse as complete as ever. Even tho the Federal Reserve Banks had made heroic efforts to stay the pervading fright, the panic burst. Gold payments were suspended, and for a short time all the commercial banks in the country were closed. All this was caused in part, no doubt, by weaknesses in the general banking structure of the United States. But it arose mostly from causes that go deeper. Private investment, private venturing, private management—the capitalistic system—seem to entail the familiar alternations of activity and depression; something of the kind seems an inevitable concomitant of progress. The periodic halt and overturn in business operations is commonly accompanied—in a way, punctuated—by the monetary crisis, easily developing into a panic. That there is an underlying connection between banking operations and these industrial fluctuations is generally admitted. The precise nature of the connection is not by any means clear. The whole broad problem of how to construct a monetary system which will at once help to prevent or allay crises and be proof against collapse is a problem which, tho more deliberately faced than in former days, is not yet solved.[1]

§ 4. The reserves which the member banks are required to hold present problems of their own. In the days of decentralized bank-

[1] On this side of the matter more is said below, in Chapter 43, Vol. II, on Overproduction and Crises.

ing all "ready money" or "cash" of any kind was regarded as reserve. Under the present system in the United States the only sum designated as reserve and required as such is that which stands to the member bank's credit at its Reserve Bank. The amount is compulsory; the law requires a given percentage to be kept at all times and under all conditions. As in the case of the Reserve Banks themselves, therefore, this "reserve," while apparently designed for use in emergency, seems nevertheless not available when the emergency arises. It must be kept intact. The appearance of anomaly is removed, however, when it is remembered that member bank reserves are thought of primarily as a means of regulating the volume of deposits. In times of panic, when need arises for cash to meet runs, a member bank which is believed to be sound can secure loans from its Reserve Bank and be credited on the latter's books for additional deposits; and against these "surplus" deposits it can draw cash and so swell its till money. This recourse to the Reserve Bank is available if that Bank thinks the member bank sound and deserving of aid in the crisis.[1]

This important element of elasticity was no part of the system as formulated in the law of 1913, but rather a tradition that gradually developed and was then confirmed by the legislation of 1935. The change is rather implicit than explicit, but it has altered entirely the significance of member bank reserves. The principal function of these reserves has become that of limiting the amount of the member bank loans and deposits; the real store of emergency cash is that available from the Federal Reserve Banks, in the form of their notes.

The member banks, to repeat, must keep in the Federal Reserve Banks deposits up to a stated percentage of their own demand obligations. These reserves are quite independent of the cash actually held in the bank's own vaults. That cash is "till

[1] The statutory requirement made in 1917 was 13%, 10%, 7% for the three classes of banks respectively. Under the discretionary power given to the Board by later legislation (the so-called Thomas amendment of 1933 and the Banking Act of 1935) for changing these figures, they came to be as given in the text—22¾, 17½, 12. Under the legislation of 1935 the percentages may not be lower than the original figures (13, 10, 7) and may not be more than twice as high. The effect of changes in them is still (1939) a matter of experiment, and it is impossible to foresee what they will come to be.

money" and each bank can hold as much or as little as it pleases. In the days of decentralized banking and of legislation adapted to it, till money was the important element in legal bank reserves. Under the new system the only sum designated as reserve, and required as such, is what stands to the member bank's credit at its Reserve Bank. The law at first blush seems to require once for all a given percentage to be kept under all conditions. As in the case of the National Bank System, the fixed minimum, apparently designed for use in emergency—a "reserve"—is yet not legally available when the emergency arises. Yet here, as in the older legislation of the United States—and also in the laws of European states which make similar requirements—the strict limit which is written in the statutes is set aside in times of stress. It serves in practice mainly as an indication of the course which the banks should ordinarily follow and on occasion may be sternly required to follow. There is flexibility both by condoning transgressions of panic days and by discretion in the use of the powers in the Board. If indeed a member bank repeatedly or habitually fails to keep its reserve up to the amount required by law, it is likely to be called on to set its house in order, but this is no part of the policy followed in times of a "run."

§ 5. Member bank deposits in a Reserve Bank may arise thru rediscounting. Rediscounting is not indeed the only process by which a member bank can add to its reserve. Others are inflows of gold coin or bullion which by law must be placed in the Reserve Banks and thence transferred into the Treasury vaults; open-market operations; and fluctuations in the Treasury balances with the Reserve Banks. But rediscounting, tho ordinarily the least important, may come to be of the first importance. It is so in case of panic,—or rather it is expected to be so in that case. Then a member bank is most likely to bring to the Reserve Bank for rediscount commercial paper which it has already discounted for its customers, the proceeds of the rediscount being placed to its credit as "reserve." The object is to obtain additional reserves, or to meet a deficiency in reserves, or to enable the withdrawal of cash. Whatever the immediate object, there is a strong tradition, amounting almost to a compulsion, that resort to rediscounting

in large amounts shall not be habitual. Among the careful and well-managed banks themselves, this attitude was taken voluntarily from the outset; perhaps in part from a mere wish on their part to be free from regimentation and control. The general policy of the Reserve Banks—intimated rather than formally announced --served to establish the same tradition against habitual rediscounting for all of the member banks. The reserve once set up, the member banks resort to it only as an occasional expedient. Free and large use of the mechanism is reserved for times of emergency. Then rediscounting causes further deposits (reserves) to arise and thereby additional cash to be made available for banks facing a run. It has already been pointed out that in 1933, when use of this procedure was made on a great scale, it did not suffice to quell a panic. Perhaps the situation at that time was the result of causes which will not recur; the episode may not be significant of what is to be expected after the amendments of 1935.[1] Yet it must be borne in mind that currency schemes have often been found to work in ways not expected. He would be rash who predicted unqualifiedly that monetary panics in the United States are a thing of the past.

It is to be remarked that the tradition against habitual rediscounting is a peculiarity of the Federal Reserve System,—of the United States. In European countries, and notably in England, the rediscounting practice is accepted as an everyday matter. The explanation of the difference lies in the development of deposit banking in the United States, its extraordinary permeation, the difficulties and disasters to which it has given rise, the need so strongly felt of guarding and regulating it. It is this aspect of the case which accounts for the way in which the requirement for member bank reserves has come to be used. The aims and procedures are of a kind hardly thought of when the system was adopted; they are the results of experience, and not unlikely to be changed again as the results of further experience. The Federal Reserve structure gradually came to be regarded, after the post-

[1] The Banking Act of 1935 greatly relaxed the legal restrictions governing the extension of credit to member banks by Federal Reserve Banks. It is now (1939) possible for Federal Reserve Banks to lend to member banks upon any acceptable security, whereas before 1935 strict rules of "eligibility" had to be satisfied.

war depression of 1921, still more after that of 1929, as a means not only of meeting runs and preventing crises but also of controlling the commercial banks, regulating the expansion of their loans and deposits, promoting stability in the monetary and credit structure, and thereby stability of the industrial structure at large. It was a new conception, emerging slowly and gradually, tentative in its application, half unconscious. Its essence was that the member banks were to be held in leash by keeping their deposits under control. The most significant evidence of the official acceptance of the wider policy was the provision in the act of 1935 by which the Federal Board could increase, within limits, the reserves which the member banks must maintain.[1] The objective in giving the Board this power was not to enable it to make banks more safe, but rather to make more effective the Board's control of member bank deposits.

§ 6. Quite different is another method of control by the Federal Reserve Banks, one long used in Europe and used in the United States also in the period following the war of 1914–18: that of changing the discount rate. It is raised when borrowing is thought to be excessive, lowered when borrowing is to be encouraged. Just under what conditions borrowing is excessive, just when in need of encouragement, is not easy to say offhand. But there are extremes in both directions which are recognizable and which suffice to justify the application of this instrument of control. How far it can be used to advantage in the earlier stages either of activity or of depression is uncertain. Apparently a high rate is more effective for checking activity than a low rate is for counteracting depression.

Still another instrument of control, expected to be used in connection with the rate of discount, is that of "open-market operations," that is, the purchase or sale of government securities in the open market by the Federal Reserve Banks. Operations of the same kind have long been carried on by the Bank of England. Dealings in government securities by a Reserve Bank (or any Central Bank) have a different effect on the monetary situation

[1] A somewhat similar power had been granted the Board two years earlier, by a provision of the so-called Thomas amendment.

from ordinary purchases and sales. When securities are sold by one person to another there is merely a transfer of money (deposits) from the buyer to the seller; the total monetary situation remains precisely as it was before. But this is not the case when such dealings are carried on by a central bank. When a central bank buys, it buys with funds in its hands, or rather with funds created by it. The deposits by which it pays for the securities result in a net addition to the existing reserve deposits. This new amount, turned over in the first instance to the seller of the securities, soon finds its way into the member banks and thence into their deposits with the Reserve Banks, thus becoming an addition to the member bank reserves. It then provides the basis for a multiple expansion of loans and deposits by them, thus making for an increase in the total supply of money. On the other hand, a sale of securities by the Reserve Banks tends to cause a contraction of money. The money given to the Reserve Banks in payment for the securities cancels an existing deposit with them, and thus causes diminution of reserves. This instrument, like the discount rate and used in connection with it, gives a measure of control over movements of the money structure.

Still a third instrument, already mentioned, was added to the armory of the Federal Reserve authorities by the Banking Act of 1935. The Board of Governors then was given power to alter, within certain limits, the reserve requirements of the member banks. This power was presumably to be used only in dealing with major problems, as in 1936 when heavy gold imports had threatened to increase excess reserves to unmanageable proportions.

There are thus, in the United States, three major instruments of control: the discount rate, open-market operations, and changes in the percentages of reserves required for member banks. The first two have been long in use. The discount rate, the most familiar, has proved less effective here than in other countries and indeed less so than tradition had led writers and legislators to expect. The instrument of open-market operations is newer and less tested by experience. Hitherto (1939) the consequences of its use has been less great than was expected. Finally, there is the

third instrument, that of changes in the reserve requirements for member banks; the latest, the most radical, apparently of great potential power, as yet quite untested by experience.

§ 7. The Federal Reserve System obviously is not literally a central bank. It is rather a centralized system, in which the member banks are the main active parts, the Reserve Banks are supporting and regulating instruments, the Reserve Board both an administrative and supervising agency and a court of last resort on questions of policy. In the twenty years from establishment in 1913 to the amendatory act of 1935, the System grew immensely in range and power. Altho in operation for less than a generation, it became in that time so far-reaching, so familiar in the banking and business world, that it was thought of as if it had always been and always would be. Yet it is still experimental, not fully tested and tried, likely to be modified. From the first it brought possibilities of great expansion of the circulating medium; as it grew and developed, the possibilities for expansion were not lessened. Those parts of the country's banking system which remained outside its direct domain—banks chartered by the States and not joining the System—were overshadowed by it. What the ultimate outcome will be is as hard to predict as is the whole course of social and economic evolution. It is tolerably certain that the trend toward complete control by the federal government over all parts of the machinery of exchange,—all deposits as well as all notes—will persist. Localistic traditions in the States may for some time still stand in the way of the restriction of their banking institutions to those investment and lending operations which are quite divorced from the monetary machinery. But in the end the entire control of that machinery by a centralized system can hardly fail to come. Just what form it will have still remains to be seen; still more does it remain uncertain just what experience it will go thru, and how successfully so mighty an instrument can be kept in control and balance.

§ 8. Some general observations may now be made about the position, functions, future of Central Banks.

In the first place, they are the accepted financial agents and depositaries of governments. That they are private corpora-

tions in form is ordinarily a mere matter of convenience. No doubt, as a matter of government bookkeeping and accounting, the convenience is great. And there are more substantial advantages. The fact that there is an organization entirely distinct from the Treasury makes it easier to discern what the government is currently doing, and imposes some check on recklessness. It probably imposes a check of a similar sort when there is urgent need of funds, especially in times of war. A Central Bank whose organization is not that of a mere government department is likely to handle the huge financial operations of war with more skill and moderation than would the government direct. None the less, all things considered, Central Banks remain creatures and servants of the government.

Second, Central Banks have a complete monopoly (or are in the way of getting one shortly) of bank notes; and these notes in turn are becoming the usual form of government paper. The monopoly position is familiar as regards direct government paper. Less commonly appreciated but more important is the fact that this paper itself is being supplanted by Central Bank notes; or rather, government paper is more and more widely taking that form. In the United States some direct government issues remain, such as the old legal tender notes ("greenbacks") left over from the nineteenth century, or the anomalous silver certificates, which are virtually government paper. But the main paper money is the Federal Reserve note, and that is in reality a government issue. In all countries Central Bank notes became after the war of 1914–18 not only the main constituent of hand-to-hand cash, but for most purposes the basis of the entire circulating medium. Even where the gold standard was nominally maintained, it was effective only for international remittances and had but slow and uncertain connection with the monetary supply and the price level. The great bulk of the world's business was done in terms of paper, and the paper was mainly in the form of Central Bank notes.

Next, all this tended to bring the Central Banks into a position of new and wider responsibility. They became the managers of the entire monetary structure. Not perhaps of the subsidiary coin; that has so automatic a character that an independent government

agency—the mint—can be left in sole charge. But both hand-to-hand cash and deposits were managed by Central Banks, perhaps subject, as in the United States, to some prescribed legislative rules but everywhere with a large margin of discretionary management.

Managed, we say. The extent to which the trend this way is an innovation can be exaggerated. In a sense money has always been managed. Coinage is a form of management; so is the bimetallic system and the gold standard. The gold standard method, tho in some ways automatic and non-managed, was modified from stage to stage. The Federal Reserve System itself at the outset had as one of its important objects a new management of the gold standard, designed both to make it stronger and to make it work more smoothly.

And yet it is true that after the war of 1914–18 a new phase was entered in which the notable writings of Mr. J. M. Keynes were both a symptom and a powerful stimulant. Management, it came to be felt, should be a matter not merely of smooth working from day to day and year to year, but of steadiness and continuity for the money mechanism over longer periods and for larger purposes. The advocacy and acceptance of greater responsibility brought a new point of view on the range of things which Central Banks might and should do. While still bankers' banks, and exerting their influence on the community at large chiefly thru the customer banks (in the United States the member banks of the Federal Reserve System), their function was conceived to be one not only to help and support these banks but to control them also. And, as experience and information accumulated, it began to be felt that control was to have as its object not only monetary steadiness and stability but industrial stability also. By monetary steadiness I mean—for the present purpose, which has little to do with paper money issues of the extreme and reckless kind—the maintenance of acceptability and solidity for the circulating medium. Monetary stability means more: that Central Banks should be not merely refuges in times of stress and panic, not merely standing on the sidelines and ready to step in when needed, but continuously active in keeping the price level stable. The other aim, industrial stability, means still more; the prevention, or at

least moderation to the utmost, of the cycles of activity and depressions and the calamitous crises. This is the aim toward which men's thoughts are turning; even tho the trend is hesitant, often unwelcome, only half conscious of the direction in which it is moving. How far and fast it will go, what paths it will follow, no one can tell; but it is unmistakable.

CHAPTER 28

GOVERNMENT PAPER MONEY

§ 1. Paper money has come to be that put out directly or indirectly by governments.—§ 2. Convertible notes.—§ 3. Inconvertible money. How its value rests on its quantity.—§ 4. Extent to which paper has displaced coin. Its convenience. Custom as affecting the proportion of paper to the total of active money.—§ 5. Inconvertible paper may operate in two ways: as hand-to-hand money, and as the basis of the entire monetary structure.—§ 6. How far government can maintain the acceptability of inconvertible notes.—§ 7. The effects of sudden large issues; extreme depreciation and collapse.—§ 8. Experiences during the war of 1914–18.—§ 9. Unexpected consequences of these experiences in countries that adhered to the gold standard.—§ 10. The specie premium.—§ 11. Problems arising when there is return to the specie standard.—§ 12. Great issues of inconvertible paper are almost invariable consequences of prolonged war.

§ 1. Two kinds of paper money are to be distinguished. First, that put out by banks in the form of promissory notes payable to bearer on demand. This was the earlier kind. Second, government money, usually similar in form to bank notes, i.e. on its face a promise to pay; but dissimilar in that no holder can compel a government to meet its obligations. While he can bring suit against a bank as against any debtor, he cannot compel the government to meet its debts. The distinction between the two kinds, long important, has come to mean less and less. Bank notes have been gradually amalgamated into the general class of government money. Most often the right to issue them has been limited to a single bank owned or controlled by the government, and the distinction between bank notes and government notes thus made quite shadowy; bank notes then are virtually government paper. Sometimes banks, while still independent and still allowed to put out notes, have been so regulated and restricted that the government in effect has stepped in as a guarantor of the notes. Such was the case with the national bank notes which played so large a part in the United States from the time of the Civil War to their final extinction in 1936. In essentials it is the case with the bank notes

issued under the Federal Reserve legislation. In all the western countries, with but few exceptions, paper money has come to be virtually government money. The obsolete or obsolescent cases of notes still issued by private banks are historically interesting, sometimes almost romantic. They have some significance for general monetary problems chiefly by way of illustration: they show in a striking way the characteristics of bank deposits and help toward an understanding of the trend of legislation on deposit banking. For these reasons and in these aspects they will receive some separate consideration in the next chapter. In the present chapter we shall say little about them and give attention chiefly to paper money put out by government.

§ 2. Tho paper money is but a portion of the total circulating medium, and tho its proportion to the whole is limited by such factors as convenience, tradition, custom, it has come to be enormously important in the monetary structures of modern times. When convertible, it still remains an important intermediate layer interposed above the specie foundation. When inconvertible it is the one foundation of the whole system.

So long as the paper note—the dollar or pound or franc or mark —is convertible, the government Treasury or a Central Bank gives gold coin or bullion for the paper at the established mint price. It is not essential or indeed common that each and every piece of paper shall be so honored. Pieces of small denominations need not be; but they are made interchangeable for large ones. There may be other items, left-overs from historical growth, which are not convertible into gold at all; such are the silver certificates in the United States. These and like items are usually received in payment of taxes and other dues to the government and thus made readily equivalent (provided the total amount of them is moderate) to the main body of the paper. This main body is usually made a legal tender for debts.

It is conceivable that in countries like the United States and England, even in France, deposits alone should be made convertible into specie; deposits, that is, in a public bank, or in banks of essentially the same character like our Federal Reserve banks. Paper money would then be treated as subsidiary coin is,—not

full legal tender, but hitched on, as such coin is, to the dominant thing. It is chiefly tradition which has brought it about that, alike in common speech and in legislation, the terms convertible and inconvertible are connected with notes alone. But the tradition is strong. It is the rarest of happenings that any one asks for gold coin when cashing a check; he asks for the familiar paper. On the other hand, curiously enough, checks are in practice used when it is gold that is wanted, even tho as a matter of law it is the paper note only which entitles the holder to the specie. The intermediate step is ignored; he who wants specie could get paper (legal tender) for his deposit; the Treasury or Central Bank is ready to meet the check in gold, usually after certification by the bank on which it is drawn.

Convertible paper in effect serves to increase the monetary supply of specie; that is, it acts precisely as if so much specie were added. That is, so much as the *net* amount added. Not the whole amount of the paper is to be regarded as additional, but only that which is over and above the stock of specie—the reserve—that may be held for redemption when asked for. The paper makes a given amount of specie serve so much farther. The analysis of its working is thus comparatively simple,—much the same as that of the working of a monetary supply composed of specie alone.

§ 3. Suppose now that the paper is inconvertible; specie cannot be got for it. Then the paper is something more than convenient hand-to-hand money, restricted in its range of operation. And it is also more than an intermediate foundation for the monetary structure, such as convertible paper is. It becomes ordinarily the one element on which the whole circulating medium rests. How satisfactorily can it perform the functions of money? What will determine its value—the price level?

The answers to these questions can be best approached by beginning with the simplest case. Assume that there is only this one kind of money; no bank deposits, no coin or specie (except perhaps subsidiary coin exchangeable for the paper). Assume too, for the present, that the paper circulates readily, everyone accepting it without question or hesitation in payment for goods. How far this last assumption is tenable is left for further consideration;

it does not hold under all circumstances. But there are plenty of cases in which the facts are in accord with it, and it is these which are most significant for the main lines of analysis.

The reasoning already set forth in Chapter 19 then is applicable: the value of the money will be governed by its quantity. If it is issued in the same quantity as the specie or the convertible paper notes previously in circulation, the range of prices will be what it was before, and the value of the paper will be exactly what that of the specie had been. If it be issued in twice the old quantity, prices will be doubled, and the value of money will be one half. These statements assume, of course, that rapidity of circulation remains the same and that the quantity of commodities and their mode of coming to market remain the same—qualifications which already have been considered. If these other factors have changed, the proper allowances must be made. The corrections do not impugn the essential principle: the value of a freely circulating paper money depends on its quantity. Tho it be quite inconvertible, tho there be no prospect of its redemption in specie, it will retain its value and perform all the functions of money. It will have one obvious advantage over specie, in that it will cost the community less. Gold and silver can be produced only with much labor; paper money costs but a trifle. A cheap and serviceable medium of exchange is substituted for a dear one. Such is the main framework of the analysis, to which all the rest is to be attached.

§ 4. We proceed now to various illustrations, explanations, qualifications.

First, something by way of illustration: the history of paper and of the light which history throws on its present and future use. European countries began the use of paper money on a considerable scale in the latter part of the eighteenth century. At its close paper substitutes for metallic money had become familiar and in everyday business passed ordinarily from person to person without question. So much did paper money become a part of the accepted order of things that the term "money" commonly meant paper money of some sort, and this was more or less consciously regarded as the predominant part of the circulating medium. And

true it is that in countries where the facilities of deposit banking are not in general use, paper is likely to bulk very large in the total of the effective circulating medium. Thus in France, where the extreme conservatism of the peasants and the considerable (even tho less marked) conservatism of the commercial and industrial classes have prevented a very rapid development of deposit banking, paper money constitutes about two-thirds of the total supply of media of exchange. In some of the other European countries, like Austria and Italy, the situation is the same. The rapidity with which in modern times paper money has supplanted specie in these countries as the everyday money of almost all people is surprising. It is one among the many signs of the speed with which the modern world is changing, and it bids fair to change radically monetary problems and monetary politics. Only in regions like the interior of China and India does specie (in the form of silver) still hold the place which it had a few generations ago thruout the world.

One main reason why paper money came into regular use and has continued to be used is its convenience. It is light in weight and small in bulk, and can be kept on hand and carried about in purse or in pocket with ease. The fact that the notes can be printed in a variety of denominations makes it possible to carry about sums large or small without trouble. The factor of convenience, however, would have no effect were it not that, partly because of the legislation governing its use and even more because of a tradition now of long standing, paper money is accepted without question by everyone. Under the law, moreover, title to the paper passes with delivery; there is none of that cumbrousness of establishing title which is usually associated with the transfer of negotiable instruments. No occasion arises—such as at times there is with checks—for one to whom the paper is offered to inquire into the integrity of the person presenting it.

It is convenience, further, that explains at once the range and the limitations of the use of paper notes. The range is wide. In the United States it includes the larger part of retail transactions. The sums involved are generally of that moderate amount for which neither checking deposits nor subsidiary coins are adapted.

As regards checks there is not only the trouble of writing checks on the spot but the hesitancy of merchants about payments in that form from persons unknown to them; from strangers payment in cash is preferred. Not less important is the convenience, nay necessity, of paper money in wages payments. Despite the growing familiarity with banks and banking operations, the great majority of wage earners do not, and indeed cannot, keep deposit accounts in banks. At most they perhaps have savings accounts. They prefer to be paid in cash of convenient form and denomination. The banks are of the same mind; it is less troublesome and less expensive to cash one large check drawn by the employer in favor of himself than to cash a large number of checks for small amounts, each of which must be handled separately by the bookkeepers and the tellers.

The part of the aggregate of transactions that is made up of the transactions in which paper money is customarily used governs its proportion in the total quantity of active money. That proportion is not so great in the United States as the ubiquity and familiarity of retail dealings and wages payments would at first thought suggest. As compared with the total of transactions, those settled by paper money are in reality but a moderate part. In France, where custom assigns a much wider use to paper, the proportion of paper in the total is large. The extent to which paper is used, embedded as it is in custom, does not readily change. True, a great cataclysm like the war of 1914–18 may lead suddenly to a much wider use of paper in place of specie. But this is merely the substitution of one form of hand-to-hand money for another. As regards any such money, whether specie or paper, the channels in which it is used change but slowly. There are no doubt fluctuations and temporary variations in the relative volume of different kinds of transactions and so in the resort to different kinds of money. In the United States, for example, the proportionate use of paper money (not necessarily or even probably the absolute use) tends to be greater in times of depression; because the recession is less in retail transactions than in wholesale and financial transactions. Again, there are seasonal fluctuations. When the crops are harvested and marketed, there is usually an increased use of paper for

"moving the crops." More is used for retail purchases at the Christmas period in English-speaking countries, at Easter on the Continent. But these are not matters of serious importance either for theory or for monetary policy. In the main the proportion of paper notes to the total of the circulating medium is determined by customs and traditions which at any given stage or period are firmly established. This settled proportion is a matter of the first importance as regards the process by which an increased volume of paper money leads to a rise in the price level.

§ 5. We pass now to the qualifications and enlargements of the simple theory. Most important are those just suggested, relating to the part which paper money plays in the monetary structure as a whole. The situation assumed as a starting point in the preceding section was of the simplest: that paper alone was the money used. The use of credit and above all the use of bank credit in the form of deposits brings about a much more complicated situation. Paper money then operates in two ways: not only as the sole medium for hand-to-hand transactions but as the basis for a great volume of bank money. During the nineteenth century the representative form of the monetary structure was that resting on a specie (gold) basis. The gold was used in the two ways just indicated; to a considerable extent in retail and other hand-to-hand transactions, and also as the basis of the bank money, the convertible paper, and the subsidiary coin. Once inconvertible paper has supplanted specie, it serves (as the specie had done) in both ways, and then its effects on prices come not merely because it is the hand-to-hand money but even more thru its powerful operation as the basis of the entire monetary structure.

No new principle is introduced by this elaboration of the analysis. There is nothing in it to impugn the fundamental proposition that the price level depends on the relation between the quantity of the entire circulating medium and the quantity of things exchanged. But the inevitable interrelation between the two uses of paper brings it about that the effects of paper on the price level are not uniform or predictable. They are not the same in different countries, or the same in any one country at different times. Until the close of the nineteenth century, the mere quantity of paper

could be treated as the preponderant element in tracing the consequences of its issue. At all events this could be done with some assurance as regards the European countries that resorted to inconvertible paper in that period, such as Austria, Russia, Italy; not with any assurance for the United States during the years of inconvertibility from 1862 to 1879. In the latter part of the nineteenth century and still more in the twentieth bank reserves and bank money came into wider use everywhere, making it necessary to be more and more cautious in explaining or predicting the effects of paper issues on the simple ground of quantity alone. Other factors also changed,—habituation to use, ease of circulation and what not. Enough has been said in this and in earlier chapters to show how the simple explanation on quantity grounds needs to be guarded; a topic on which still more will be said as we go on.

It is somewhat surprising, in view of all the complications, that during the war of 1914–18 monetary history showed a marked accord with the simple hypothesis. The general price level under paper money régimes showed a closer relation to its quantity than might have been expected. This was at all events the case where the paper was put out in what one might call a decent or reasonable way; where the amounts were not so huge as to bring about complete collapse, or by decrepit governments whose financial and monetary future was desperate. England, France, and (during the war itself) Germany all resorted to paper, and succeeded in maintaining their internal prestige and credit sufficiently to keep the paper in free and steady circulation. A great rise in price ensued in all; and that rise was roughly in proportion to the increase in the quantity of paper. It is to be said that the same kind of approximate proportionality appeared in earlier cases where there was depreciation but not collapse; in European issues of the nineteenth century, and even in the United States issues of that century.

§ 6. We pass now to some questions which are more peculiar to paper money. What has been said in the preceding sections is hardly more than an application of conclusions which hold of any kind of circulating medium. The question peculiar to paper arises chiefly with regard to the extent to which the government can uphold and regulate its value,—its purchasing power.

Inconvertible paper has often been called fiat money, the implication being that its circulation and value are settled solely and definitively by the edict of the sovereign or legislature. The implication is a dubious one. The extent to which the mere command of the political authority can cause a piece of printed paper to serve as money and to maintain its value can be both overstated and understated. Historically, all money has had its origin, directly or indirectly, not in any compulsion or even in any deliberate selection, but in the customary acceptance of some commodity of general serviceability. But when a commodity has once come habitually to be used as money, public authority can very much affect its value and the mode in which it circulates. So is it with paper notes. Only when people have become accustomed to using notes as a medium of exchange can paper pieces be made to serve as money by mere government fiat. Given the established custom, however, it is quite possible, as governments early discovered, to make paper notes circulate freely whether or not the printed promises to pay (that is, to pay specie) were kept.

Even when the custom does exist, there are limitations. Fiat money does not necessarily circulate with freedom merely because it bears a government stamp or promise. Conceivably people will distrust the government, or dislike to use paper, or for one reason or another refuse to accept it readily in current transactions. Of this possibility a striking illustration appeared in the state of California during and after our Civil War, from 1862 to 1879. Paper notes not convertible into specie, issued at Washington by the government of the United States, circulated freely everywhere thruout the North except in California. People in that state, tho they neither distrusted the government nor felt hostile to it, had a strong feeling in favor of gold and against paper. They resolutely conducted their transactions in gold, tho all the rest of the United States used the inconvertible notes. But the Californian case is exceptional. When those primary conditions are satisfied which are essential for the circulation of paper notes—habituation to the use of paper and a firm government—the paper passes readily. Almost always too the government makes the paper a legal tender for debts and also receives it for taxes and other public dues; thus

exerting its full strength to bolster up the fiat money and make it circulate freely.

How long general acceptability will last, whether and when collapse will come, depends partly on the amount issued and partly on political, military, psychological factors which are not easily gauged. Moderate increases in volume, say of 50 per cent or even 100 per cent, may rouse no uneasiness, and ready circulation may well continue. Even a doubling or quadrupling may not cause serious disturbance, if it takes place slowly and quietly. So much was shown by the experiences of England, France, Italy during the war of 1914–18. But when the increase is sudden as well as great, and the price level rises dramatically, a new factor enters—lack of confidence. Then disruption threatens. A situation of this kind is most likely to be engendered when the public finances are subjected to sudden and heavy pressure. In time of war, the drain on the public treasury may abruptly reach immense proportions. A government which is faced with huge expenditures and has exhausted its powers of taxation and of borrowing, feels compelled as a last resort to print once for all the money needed to pay its mounting bills.

§ 7. What then happens? Suppose a government, hard pressed for funds, prints money in great and rapidly increasing quantity. Its value declines steadily and quickly. If the process is carried far, an increase in velocity sets in. As people see the value of the money go down even in the brief period during which it is in their hands, they make haste to pass it on quickly for goods. The value of the paper then falls more than in proportion to the increase of its quantity. If the increase of issue still persists, the stage may be reached when no one will accept the paper. The bottom drops out completely, very likely with dramatic suddenness. Its value disappears not merely because both quantity and velocity are very great, but also because people are no longer willing to accept it at all. The demand for money (the offer of goods) may cease entirely. Such is said to have been the case with the notes which the Scotch schemer and adventurer, Law, persuaded the French government to issue in 1720. They were put forth in such enormous and increasing amounts that they completely lost acceptability

and depreciated to nothing.[1] Such was the case, too, with the paper money issued by the American Congress during the Revolution. "Continental" money was printed in amounts so vast that it became utterly distrusted, and depreciated much more than in proportion to quantity (whence the saying "not worth a Continental"). Such, too, was the case with the assignats of the French Revolution in 1790–96, when the French government put out notes which at first were redeemable in land, but soon were poured forth without pretense of any redemption, and in such unlimited quantities that they became quite worthless. Still later, in 1864–65, the same was the fate of the paper money of the Southern Confederacy.

§ 8. The various stages in the use of paper and in its depreciation appeared strikingly during the war of 1914–18 and the years of distress and confusion which followed the peace. All the warring countries (with the sole exceptions of the United States and Japan) had recourse to this tempting way of "raising money." In some countries, tho there was depreciation of value and a marked rise in prices, the paper functioned normally; that is, it circulated freely, sales and purchases of goods being made as in ordinary times. Such was the case in Great Britain, France, even in Germany before the surrender. Some exaltation of patriotic fervor doubtless played a part in preventing anything like a collapse. It is cases of this kind, where prices rise markedly but business goes on as before, that bring out most clearly the strictly economic aspects.

In other countries the issues were so huge in amount—in Russia, Austria, Hungary, Poland, the Balkan countries—that the last

[1] The breakdown of confidence in the paper seems to have taken place in this case with dramatic suddenness. An effort by the government to put a limit to depreciation caused an unexpected and utter collapse. During the first stages of depreciation, "strange as it may appear, the deterioration of the notes in value does not appear to have affected their circulation. All that people looked to was nominal value, and while the notes were called livres, nobody inquired what a livre meant. But the instant the denomination was altered; the instant the government declared that a note for ten livres should be worth only five,—the baselessness of the paper fabric was detected. The terror was as universal and as blind as the confidence had been. To use Sir James Steuart's words, on the 22d of May, a man with one hundred millions of bank notes might have starved in the street." Senior, *Three Lectures on the Cost of Obtaining Money*, p. 76. The reference is to Sir James Steuart's *Principles of Political Economy*, Part II, Chapter 59 (Vol. III, p. 52, edition of 1770).

and catastrophic stage was reached. Most striking of all was the course of events in Germany. The German paper issues of 1923 constitute one of the most extraordinary episodes in monetary history. A country in which the level of intelligence was as high as had ever been known, whose previous monetary policy had been marked by care and restraint, whose financial leaders and advisers were the same as in the preceding years, plunged into paper inflation to an extent nothing less than fantastic, and with results even more far-reaching and calamitous than in any of the earlier experiences. Just how it happened has not yet (1937) been uncovered by the historians, but the main happenings are easily told. There had been great issues and heavy depreciation even before the occupation of the Ruhr in 1923. That ill-starred step not unnaturally brought the German government, already sore pressed, to a feeling of sheer helplessness; it was ready to turn to any device, even the most dangerous and discredited, that promised monetary relief. Yet it is a mystery why paper money was issued in such preposterous quantities,—not only in terms of millions of marks, but of billions, trillions, quadrillions. The entire machinery of exchange was disrupted, and at the last stage the nominal circulating medium became nothing more than bulky bundles of worthless printed paper. Debtors paid creditors in paper of this kind. A sober, intelligent, hard-working middle class was almost completely dispossessed, and feelings of helplessness and resentment were engendered which deeply affected not only the subsequent course of political events but the very foundations of the social and economic structure.

§ 9. Nothing in all this was novel, except that never before had the resort to inconvertible paper been so ubiquitous, and never before had the evil results been experienced on a scale so vast. The consequences as regards the use of paper and the course of prices were in the main such as could have been readily predicted. A further consequence ensued, however, not indeed out of accord with the reasoning accepted by economists, but unexpected. The extraordinary resort to paper brought great and rapid price movements not only in the countries where the paper itself was issued, but in others which still maintained the gold standard. The range

of prices—the value of gold—was profoundly affected in the gold-using countries themselves, and the gold standard itself came to be endangered in all.

In a previous chapter it has been pointed out that changes in the total supply of gold take place slowly. This is because the total volume of gold is so great that no increment can be added in a short period which will materially change the total monetary stock. Stability in the value of money is thus supposed to be secured, at least in so far as the quantity of gold affects it. And this stability is supposed to be preserved for each and every gold-using country, since if there be a sharp increase in the gold supply of any one, tending to cause within it expansion of money and a rise of prices, the currents of international trade will carry off most of the increase of gold and hence prevent any rapid advance in prices in the country first affected. Suppose, however, that half the world suddenly gives up gold, resorting to inconvertible paper; and that the gold previously used in this half is sent to the other half, accumulates there, cannot overflow. Then the conditions on which stability depends are seriously affected. A rapid advance in prices may ensue in countries which continue to use gold.

This is substantially what happened in 1914–18. Half the world or more—almost all the warring countries—resorted to paper, and a large part of the gold held by them was driven abroad by the advance of prices. The gold made its way to the rest of the world,—the neutral countries, such as Holland, the Scandinavian nations, Spain, and to the United States and Japan. It is true that the process by which the gold became congested in these areas was not so simple as would ordinarily be assumed in economic reasoning. By no means all the gold flowed out of Great Britain, Germany, France; much was impounded at home by governments. Nor did international trade take the precise course analyzed by economists for the conditions of peaceful trade. But these divergencies from the course which trade and industry follow in times of peace do not affect the phase of the situation with which we are here concerned. Gold in great quantities was sent to a small number of countries, caused rapid and great increase of the quantity of money in all of these, and brought about a revolution

in general prices of a kind which had been thought to be so improbable from a change in the gold supply as not to be of moment in weighing the advantages and disadvantages of the gold standard.

The United States felt this influence not less than the other regions affected, the more conspicuously because of the country's great resources and their importance in the ebb and flow of the war. During the years 1915–17, when the United States was still neutral, a round billion of dollars of gold came into the country from the warring nations. The supply of gold was nearly doubled. There followed an increase of bank notes and other paper money, an expansion of bank deposits, and a rise in prices. The price rise could not have taken place, certainly could not have been maintained, but for the fact that there was no large area to which an overflow of the metal could take place.

Hence the extraordinary advance in prices which took place between 1916 and 1919. Wholesale prices nearly doubled, a rise no less great and no less rapid than that experienced during the Civil War régime of inconvertible paper. The same results ensued in the neutral countries which continued to use gold, and in Japan also, which enjoyed with the United States the dubious distinction of having entered the war and yet maintained the gold standard. The entire world, warring and neutral, specie basis or paper, was involved in a monetary revolution the like of which, in range and intensity, was unprecedented in history. All the evil consequences ensued, familiar in kind but beyond example in degree. The relations of debtor and creditor were turned topsy-turvy. While in some countries, notably in Great Britain and the United States, wages rose probably in quicker correspondence with prices than on previous occasions of inflation, they did not in general rise as fast as prices, and a multitude of business enterprises reaped profits—sometimes fabulous profits. Persons whose money incomes were held fixed by the inertia of custom or contract, such as clerks, manual laborers not directly reached by the main current of inflation, teachers and other public employees, suffered severe distress. Side by side came accusations of extravagance and complaints of rising expenses; the natural consequences of the fact

that in some cases money incomes rose more than in proportion to prices, in others less. Notwithstanding the fact that the gold standard, the instrument which is supposed to have the one decisive merit of stability, was unimpaired, the general range of prices and incomes and the relation between individual incomes and prices were utterly unstable. And if this was the case in the countries which held to the gold standard, how much more in those resorting to paper! Never had there been thruout the world such monetary chaos.

The chaos did not cease with the war. The worst of the paper money excesses were got rid of by the middle of the decade 1920–30. But the underlying economic distortions were by no means smoothed out, were among the causes that led to the crisis of 1929, and made immensely more difficult the process of recovery from that unexampled collapse.[1]

§ 10. The specie premium which appears in a country resorting to inconvertible paper is familiar. Tho in some ways obvious, its significance is often misunderstood. It is supposed to be a clear proof of depreciation and a certain measure of the extent of depreciation. The matter is not so simple.

Specie payments are seldom abandoned until they can no longer be maintained. What happens is that almost from the first the expansion of money tends to drive out specie. The expulsion takes place thru the operations of international trade. The increased quantity of money in circulation sooner or later raises prices. The rise in prices causes imports to be greater, exports to be less; and specie flows out in payment of the imports. Paper money, of course, does not flow out; it cannot circulate in foreign countries. The usual mechanism is not described with accuracy in all this; sundry complications in its working will appear when the subject of foreign trade is given detailed consideration. But the essentials of the process are as here stated. Specie disappears thru the channels of international trade, in proportion as the money supply is increased and prices raised.

Gold never disappears entirely. It does indeed disappear from

[1] On the events of this period and the questions they raise about the gold standard, see Chapter 30, below.

circulation; in the monetary system it is supplanted by paper. But all gold does not necessarily leave the country. Some is always wanted for use in the arts; and for these uses it is bought and sold, like copper or nickel. Some is commonly wanted also for transactions which by special stipulation are to be carried out in gold. A class of dealers in gold usually appears, who make it a business to buy and sell the metal, as other dealers do with the more common metals. The only peculiarity is in the way the price of gold is quoted; it is often stated, not in terms of so much per ounce or pound, but in terms of how many paper "dollars" are needed to buy one gold "dollar." What arises is a "premium" on gold. Aside from this the metal, being no longer a money metal, becomes like any other commodity.

The premium on gold roughly measures the depreciation of the paper, but measures it no more than roughly. If the premium is steadily rising, it may be inferred that prices are rising in roughly the same degree above what gold prices would be. The broad movements of the premium give an indication of the general change in the price level. When for example the premium, after having been high for a period of years begins steadily to fall, we may infer that paper prices are coming nearer to what gold prices would have been—that they either are falling, or are failing to rise as gold prices elsewhere are rising. The actual depreciation of the paper, however, is the rise in prices. This we try to measure by the index-number method. But any rise in prices, as has been seen, is irregular. Some commodities advance more than others, some not at all, some decline. The change in any one may or may not be such as to indicate the general change. So it is with the price of gold, or the specie premium. It is subject to influences of its own, among the most important of which is the demand for remittances abroad—the use of gold as a standard for transactions with foreign countries; not only for imports and exports, but for the so-called invisible items. The complexity of the foreign trade operations brings it about that sometimes the premium is in advance of the general rise in prices, sometimes lags behind.

One of the factors which lead to special fluctuations in the gold premium is the prospect of the redemption of the paper in gold.

Any event which makes early redemption in specie probable lowers the premium; any untoward event raises it. When Napoleon broke loose from Elba in 1814, the premium on gold in England rose; when the news of Waterloo came, it fell sharply. In the United States, during the period of the "greenbacks" in the Civil War, the premium fell at once after the battle of Gettysburg and rose high during the anxious summer of 1864. Such abrupt fluctuations have led to the statement that it is confidence in the paper money which governs its value once for all, or at least mainly affects its value. It is more accurate to say that confidence in redemption affects the value not of the paper but of the specie. General prices do not move up and down under the influence of military or political fortunes. Dealers and speculators in gold and in foreign exchange discount at once the consequences of a victory or an election for the financial stability of the government and for the possible resumption of specie payments. It is in this way that political events and prospects are likely to affect the price of specie in terms of paper.

§ 11. When a return to specie payments is to be made, a difficult question presents itself. Shall the paper be redeemed at its face value or at its market value? The first course has the bracing effect of recognizing a promise to pay as being really a promise, and of meeting it to the letter. Yet in many cases, perhaps in most, it is doubtful whether such a course is desirable. If the money has been depreciated for a considerable time, a return to the old specie parity offers no great advantages and carries with it serious disadvantages. Business contracts have been made on an existing scale of prices and money incomes. Debtors and creditors alike have made their engagements on that basis. To shift these into specie engagements, and to enforce them during and after the transition to prices presumably lower, is to injure the present set of debtors as much as creditors had been injured in the past. Moreover the process of reducing the money supply and forcing prices to a lower level is very likely to lead to a stage of business depression.

Hence the best plan may be to "devalue" the money. The gold standard can be restored and the paper made convertible into gold; but a smaller amount of gold may be put into the standard

unit than formerly. If the money is thus depreciated to, say, only a half of its former value, the plan would be carried out by redeeming each unit of money in only half of its former gold content. A method of this kind was used by several of the European nations when returning to gold after the war of 1914–18. In France, for example, the new franc was made the equivalent, in gold, of about a sixth of the old franc. This plan is the best when the depreciation has lasted for a considerable period, and there is the more reason for adopting it if the depreciation has been severe. The greater has been the rise of prices, the greater must be the fall in order to restore the old position; and the larger the required fall of prices, the more troublesome and long drawn out are the resulting dislocations. If the depreciation has been both severe and long-lived, there is hardly a choice: it is best to devalue boldly.

But where the depreciation is not of long standing; where return to a specie standard has been steadily expected, and has been borne in mind as at least a possibility by all lenders and borrowers —there the sound policy is to resume at par. Especially is this the case if the degree of depreciation has not been very great. Redeem the money at its full nominal value, and maintain the good tradition that "a dollar is a dollar." Doubtless it is a half-illusory tradition. The doctrine that it is honest to redeem a paper dollar in gold at its face value implies more as to the nature of "honesty" than the average man will understand, but it is not to be caviled at unless there be very serious grounds for questioning its substantial equity.

At all events, the return to specie payments has often taken place by resumption at par. So England proceeded after the Napoleonic wars. Such was the plan followed in the United States in 1879.

Great Britain made a brave attempt—perhaps in the light of later events an ill-judged one—to do the same in 1925, when specie payments were resumed at the pre-war par. The attempt, continued for five years, was given up in 1931. The causes of that failure were complex. Partly they were of a primarily economic kind; perhaps not so much the troubles from readjusting the debtor-creditor relations as the depressing effect on trade and industry from a widespread fall in prices. More important were

obstacles of a semi-political kind: the strength of the labor vote, the obstinate refusal of the labor unions to accept any decline in money wages, the special importance in the country's economy of the export industries in which the labor unions were strongest. The "stickiness" of money wages (to which reference has already been made) was reinforced not only by the militant trade unions but also by the practice of government regulation of wages which the war had so greatly promoted. The depression which set in after the crash of 1929 gave the final blow; and specie payments were given up in 1931. That momentous overturn profoundly influenced monetary policy the world over. More than any other one event it opened wide a new set of questions: whether the gold standard had ever really worked well and whether any specie standard should be maintained at all. These are crucial matters in monetary theory and practice, which cannot be well handled until still other phases of monetary theory and history have been taken up and hence are reserved for later discussion.

§ 12. The immediate cause of resort to inconvertible paper has almost always been the pressure of public finance. A government that is in urgent need of funds with which to meet its expenditures finds increased taxation too difficult or too unpopular, is reluctant or unable to borrow in the ordinary manner and turns to the printing press. Resort to this easy way of meeting public expenditures has usually been the consequence of war. Almost all the well-known cases of notes utterly discredited—the assignats of the French Revolution, the Continental money of our own War of Independence, the Confederate notes of 1862–65—all arose from the stress of war. Other issues which reached the stage of depreciation, tho not of complete collapse, were the consequences of the same stress. During the Napoleonic wars England resorted to paper money (in the form of Bank of England notes, made inconvertible by law). Prussia, during the same period, turned to direct state issues. In Austria the wars of 1853, 1859, 1866 brought about a persisting paper money régime during the greater part of the nineteenth century. Russia's experience was very much the same. The United States had her trying experience during and after the Civil War and would have had the same experience during the war

of 1812–15 had that war lasted a few months longer. If the United States had entered the war of 1914–18 at an earlier stage, and if that war had lasted much longer than it did, it is probable that then too the country would have in large measure financed the huge expenditures by resort to paper.

These experiences must be borne in mind when considering—as presently we shall—projects for radical change in the monetary structure. Usually these involve, in the last analysis, the use of some form or other of money created by the state and managed by the state. Political policy and military needs then become the sole determinants of monetary legislation. No doubt these factors become of enormous importance thruout the economic sphere; government, international policy, and economics are inextricably interwoven. But the connection is especially close in the case of money. If there be another war like that of 1914–18, we may be sure that the monetary experiences and distresses of that period will be repeated. No wish or effort for restraint will prevent the resort to paper, and to increasing resort as war lasts longer and becomes more intense and frightful. No international agreement, no new kind of monetary standard, no ingenuity of the ablest statesmen and the best economic advisers will prevent the old familiar excesses. Victory will spell monetary calamity, defeat monetary disaster. Even in times of peace, monetary policy is difficult, troublesome, beset by economic uncertainties and political passions and dangers. In war any and every good system of money and finance may be overwhelmed by an irresistible flood of inconvertible paper. To such dangers has the world been brought by the gradually increasing use of paper. All the conveniences which it has brought, and which, with well-guided and careful management, it might conceivably bring even more in the future, may then prove to be outweighed by its enormous potentialities for harm.

CHAPTER 29

THE THEORY OF PRICES IN REVIEW

§ 1. The Quantity theory once more. The term "money." How far gold remains "money."—§ 2. Barter transactions, and others settled without use of cash. Credit where there is mere postponement of payment. "Bank credit," as this term is often used, raises different questions.—§ 3. How far the two factors—volume of money and volume of goods—are independent. Effects of changes in price on volume of production.—§ 4. Effects of changes in production on prices. Deposits when created in conjunction with changes in transactions. Acceleration of prices in the upswing period of a business cycle.—§ 5. The bank rate of interest and the price level.—§ 6. The case of government paper.—§ 7. What remains of the Quantity theory?

§ 1. We return now to the vital topic of monetary theory; the relation of the quantity of money to prices, and the way in which the general price level is determined. In an earlier chapter of the present book (Chapter 19) it was argued that under the simplest conditions—money a physical thing, its rapidity of circulation constant, and the things in process of exchange a given physical quantity—prices would vary exactly with the quantity of money. A supposition of that sort underlay the reasonings and explanations of the several succeeding chapters. And it is to be said that, even tho an exact formulation (e.g. "the elasticity of demand for money is unity") is no more than an introductory statement of the bare elements of the problem, the general proposition that prices go up and down with the increase and decrease of the money supply continues to underlie the discussions alike of economists and of laymen. We may now try to draw the threads together and indicate what may still be the meaning and significance of the "quantity theory."

At the outset, recall the meaning of the term "money" as used in the preceding pages. It means whatever passes readily from hand to hand in settlement of transactions. It includes the coins—subsidiary coins as well as those of full-value specie—which are in everyday use by the public; it includes the paper money, whether

issued by the government or by banks, which is in ready circulation; and it includes bank deposits which are readily transferable by means of checks. Money includes, in short, whatever is in fact generally used as a medium of exchange.

Not all of the things which might be thought of as money will meet the requirements of this definition. "Savings deposits" in banks, for example, tho often regarded by their owners as the equivalent of money and sometimes in fact used as a convenient temporary lodgment for available income, are not the stuff itself.[1] They cannot be passed from person to person in the way demand deposits are by check. Finally, to cite a unique and somewhat paradoxical example, gold itself, under the circumstances prevailing in the United States at present (1939), is not money in the sense in which that term is used here. Despite the fact that our monetary system is in a way established on a gold basis, and despite the fact that gold itself is to be had for use in international payments, the gold cannot be counted as money for the reason that neither by law nor custom can it be used within the country as a medium of exchange.

This last example suggests a matter which has not yet been given the emphasis it deserves. It is that in measuring the quantity of money in its relation to prices attention should be fastened on the effective supply,—the various media of exchange that are actually in the hands of the public. Otherwise there is danger of including something which, tho it may be in appearance and in physical substance precisely the same as the pieces passing from hand to hand, has not in fact the same effect because it is not directly available for use. Coins and paper notes in people's pockets and in merchants' tills, ordinary bank deposits subject to check,—these obviously must be included. Cash held by banks in their reserves, or deposits held as reserves in central banks, are not to be included. These may be drawn upon, to be sure, and when they are, the funds so drawn move into circulation. But it is better not to count them as part of the money supply until they have been put into

[1] Exception might be made of time deposits. Tho not subject to check in the ordinary manner, these are sometimes used by business men of the United States in making payments. The practice is not general, however, and no violence is done the facts in saying that broadly checking deposits alone are to be counted as money.

circulation. Then only do they begin to serve as a medium of exchange. Hence it is that in the United States at the time these lines are written, gold should be excluded, there being indeed a legal prohibition which prevents the use of gold for actual payments. By the same token, we exclude once for all any notes which have been printed but not yet issued, coins which have been manufactured but are still in the mint, and promissory notes or other documents which may be the raw material for bank deposits.

§ 2. The general principle may first be restated summarily. Other things equal, the general level of prices varies directly in proportion to changes in the quantity of money. An increase in the quantity of money tends to bring about a corresponding rise in prices; a decrease in the quantity of money tends to cause a corresponding fall in prices. Money is a medium of exchange, a counter. Given, then, the amount of commodities to be exchanged and the number of times each piece of money is used, any alteration in the total quantity of money can result only in an alteration in the number of pieces of money used in each transaction. The fundamental relation of money to prices thus is derived from the purpose for which money is used.

This proposition, however, rests upon certain suppositions of which it is necessary to take account in applying the theory to the complex real phenomena. Some of these suppositions have already been noted,[1] and need only to be elaborated more fully; others remain to be stated.

First, the simpler qualifications and explanations. The immediate relation of money to price holds only for those transactions which are settled at once by the payment of money. Not all transactions are thus settled. Some involve no use of money at all. There are relics of dealings which are in the nature of barter. In rural areas of the United States farmers still dispose of their produce thru the "country store." In parts of New England eggs and garden vegetables are still regularly received by village storekeepers from the country folk and credited against their purchases. A farmer has an account with the store, upon which "credit" is given him for whatever produce he brings in and from which is

[1] See Chapter 19, above.

deducted the charge for what he buys. In large part the credits and the debits offset; only the balance is paid off periodically in money. Such dealings—i.e., those which offset—present no difficulty in the theory of the relation of money to prices. There is merely a barter system, limited in range, which happens to persist in a community most of whose exchanges take place thru the use of money. The barter area is so small as to be negligible, and in it the barter takes place in terms of prices accepted from the dominant monetary sphere.

Another sort of qualification, less simple but of greater substantive importance, is to be made when full payment in money is expected and agreed upon, but is postponed for some considerable time. Purchases often are made on "book credit," not only in retail transactions but in wholesale. The buyer gets the goods at once, while the sum called for in payment is set down as a charge against him on the books of the seller. The payment, instead of being made at once, is postponed. Such purchase on credit has at the time the same effect on prices as immediate payment. If in addition to a given number of purchasers who offer money there are as many more who offer to buy on time and who are fully trusted, the effect on the seller is the same as if the entire number offered money. Given the same supply of goods, prices would double in either case.

But this is the effect only during the first stage. Sooner or later the goods bought on credit must be paid for. In the end, when the debt comes to be paid, money will be used. More than this, the money which is so used will not be available for other sorts of transactions. To the extent that money is dispensed with at the outset and so set free for other transactions, to that same extent more money is called for the subsequent repayments and hence at that later time less is available for other transactions. The rise of prices which is caused at the outset has its counterpart in a fall of prices at the end.

For a period, then, the extension of book credit may have the same effect on prices as an equal increase in the quantity of money, and the eventual discharge of the debts the same effect as a decrease. There is no certainty that quantitatively the two sets of

operations will offset each other with exactness. But on the whole there is a rough equivalence. In the great pendulum swings of modern industry there are apt to be intervals of considerable length—a year or two, perhaps more—when new purchases on credit are made more freely than repayments on account of earlier purchases. During such a stage credit operations act to raise prices by an amount corresponding to the net increase of credit. In time of recession the situation is the reverse,—the payments of old debts exceed the new purchases on credit. Shifts like these, tho probably not of great consequence, play some part in bringing about the oscillating movements of prices.

Dealings of both these kinds are usually described in business parlance, and rightly, as "credit" transactions. This use of the term must be distinguished from its use when "bank credit" is spoken of. The dealings just described are a means of postponing payment, whereas bank credit is a means of making payment. The one affects the use of money, the other—bank credit in the form of deposits or bank notes—*is* money. The promise of an individual, even tho it be in the form of a promissory note, can hardly circulate freely, however well known and reputed he may be; if only for the reason that it must be a matter of accident whether it be of a denomination convenient for another transaction. No doubt it is possible that the receiver of a promissory note may endorse it and turn it over to another person in payment of goods, thus causing the note to circulate for a time. This kind of circulation of promissory notes of individuals seems to have taken place on a considerable scale in England in the first part of the nineteenth century. But it is hardly more than one of the many curiosities of monetary history. "Credit" in the form of bank notes and bank deposits, on the other hand, has played a great and growing part in the means of payment. The notes are made out in denominations convenient for hand to hand use; the deposits are made even more convenient and effective by the mechanism of checks and clearing.

It may be remarked at this point that the term "bank credit" is frequently used in still a different sense. Like all expressions adopted into economics from ordinary language, it has no single precise meaning. Quite as often it refers to the loans made by

banks. An increase in the total volume of loans made by banks is spoken of as an increase of bank credit; and the same language is used about an increase in the total volume of bank deposits. The fact that the two commonly vary together leads them to be spoken of as if they were the same thing; whereas, tho connected, they are not at all the same or always connected in the same way or to the same extent. In view of the ambiguities of the phrase, which conceal some very difficult problems, it has been the endeavor to refrain from its use in this book. Yet it is extremely difficult to be rigorously consistent in matters of this kind. The use of familiar terms in different senses is so imbedded that varying meanings slip in quite inadvertently. On every subject in every language the reader must sometimes infer from the context just what is meant by a particular word or phrase.

§ 3. The preceding sections have dealt with qualifications and explanations which can be fitted readily into the logic of the quantity theory. We have now to consider something different,—modifications which are important as a matter of reasoning and no less important as regards the facts to which they apply.

Recall the essential content of the theory. It states a direct proportional relation between the amount of money and the level of prices, the velocity of circulation of money and the quantity of goods being assumed to remain unchanged. In the simpler formulation of the theory, the quantity of money, velocity, and the volume of goods are treated as independent variables. Each factor is regarded as being governed by causes apart from those bearing on the other factors.

In good part, however, the several elements are not thus independent; they are interrelated. When one changes, the others may respond. An increase of the money supply, for example, does not merely raise prices; it may also cause an increase in the amount of goods produced and marketed; and usually it is accompanied by a rise in the velocity of money. A reduction in the volume of money may bring a diminution in the quantity of goods. Nor is the connection only in the one direction, merely from the money side to the goods side; it runs also the other way.

The effects of changes in the monetary supply on production

are chiefly of the short-period kind. In the long run and on a larger view, the volume of production depends not on prices or the price level, but on the state of the arts, the natural resources, the stuff of the men and of their industrial leaders. The short-run influences of monetary change on production show themselves most markedly in times of rapid paper money expansion or contraction, and even more familiarly in the successive phases of the business cycle. Such periods, tho short in a nation's history, may run for several years, and each while it lasts is likely to be regarded by most people (business men included) as if its conditions would persist indefinitely.

On the other hand, there is an interrelation from goods to money; changes in the volume of production lead to changes in the quantity and velocity of money. This is in the main a long-run matter; complicated, it is true, by short-run variations, but constituting a permanent element in the working of the monetary mechanism of deposit-using countries. How far can these facts be fitted into a quantity theory?

Consider, first, the volume of goods exchanged, and the reasons why this volume changes along with the quantity of money. The explanation lies in the uneven responsiveness of prices. As was explained in an earlier chapter, individual prices and incomes do not move at the same pace in the course of a general movement. Partly because of the contractual nature of many costs, partly because of the fact that production is spread over time, and partly because of sluggishness in the workings of competition, the prices of some commodities, and more particularly the wages of labor, lag behind the broad drift. Hence a movement of prices is usually accompanied by a disturbance in the structure of the prices of finished goods and their costs of production. On the rise, some costs—notably wages—increase but slowly, and as a result profits are swollen; on the fall, since the costs come down with relative sluggishness, profits are shrunk. Since production in general moves with profits, an increase in the quantity of money, tending to raise prices, draws forth increased production, while a decrease in the amount of money, tending to lower prices, occasions a curbing of production.

No doubt some effect of this kind comes also from the mere

psychological effect of rising prices. Even tho all money incomes and all prices of goods might rise together and to the same degree, people would *feel* more prosperous, accustomed as they are to gauging all prosperity in money terms. It is only after an interval that they will begin to reflect that even tho they have more money, they are not in substance better off. Perhaps—who can say?—the illusion would persist if an upward movement went on indefinitely; not indeed if the pace were the reckless one of the familiar paper money excesses, but a moderate pace, never relaxed yet ever progressing. Certain it is that this psychological effect has sometimes continued over periods of considerable length. During the last two centuries there have been decades and longer periods during which prices clearly moved upward, others when a downward trend was unmistakable. During the upward stage, people not only feel better but go ahead more confidently; the reverse during the downward stage. Production responds. Rising prices give a zip in industry, which may indeed develop into the speculative stage but for the time being brings not only cheerful feelings but increase of the material output.

§ 4. Turn now to sequence of cause and effect in the opposite direction. A change in the quantity of goods and in the volume of transactions may bring with it a change in the total quantity of money. This is the case most strikingly and on the largest scale with deposits. A purchase of goods, the discount of commercial paper, the creation of deposits—all these go together. An increase in purchases leads to greater borrowing and a larger volume of deposits. The increase in transactions leads in itself to an increase in money. Likewise a decline of transactions means less discounting at the banks and a smaller volume of money. Nor is this true only in the case of deposits. The same interrelation obtains when notes are issued with freedom and elasticity. Then a part of the proceeds of loans is taken from the banks in the form of notes instead of in deposits, and more notes are issued in response to a greater supply of commodities. The issues of the Bank of France and of the Canadian banks fluctuate from week to week as more or less transactions are to be effected.

This kind of connection between the volume of goods ex-

changed and the amount of money is sometimes described by saying that, where transactions are implemented by borrowed funds, the goods "create" their own medium of exchange. In a sense this is true. Merchants and manufacturers borrow from banks in larger or smaller amounts according as their requirements for the purchase of goods and services are greater or less. Moreover, the size of their requirements depends not only upon the physical volume of their purchases, such as the number of tons of coal to be bought or the number of men to be hired; it depends also upon the valuation of these things,—the price of coal per ton or the wages of labor per hour.

Yet it is better to avoid speaking of *transactions* as "creating" money. In any strict sense, the statement is obviously not true; interpreted loosely—which it must be to have any meaning at all —it is misleading. Sometimes it has been argued that the quantity theory statement is just opposite from the true one; it says that prices govern the quantity of money, whereas the relation in fact is just the reverse. Those who believe this reverse relation to exist draw the conclusion that regulation of the amount of bank money is quite unnecessary; from the nature of the way in which this money is created it cannot possibly exceed the transactions which it is to conduct. The thing necessary, it is said, is to confine the creation of money thru loans to actual commercial operations; nothing more is needed.

This view cannot stand the test either of logic or experience. That prices move up and down, even when the creation of money is confined to ordinary commercial processes, is familiar. That these price movements are connected with changes in the supply of money is not open to question. Nor is this different from what the logic of the matter would lead us to expect. As we have seen, deposits when created, whether for business or other purposes, remain in existence for some time. Even after the completion of the particular exchange for which money is created by a loan, the deposit continues. It becomes a part of the sum total of deposits and is passed from person to person by check. It may be so used perhaps a dozen or more times to effect as many additional transactions after the one for which it was created. In relation to these trans-

actions, the money constitutes an added quantity of purchasing power not matched by any added production.

Yet it remains true—to return to the main line of reasoning—that the supply of deposit money is not independent of the exchanges to be made. Greater activity of business and a higher level of prices tend in themselves to bring about an increase in the deposit medium of exchange. This does not warrant the assertion that the amount of such money is determined with any exactness by the volume of transactions; but it points to an interrelation which exists and which, even tho its quantitative bearing is hard to measure with precision, cannot be set aside as having little weight in the modern world.

Turn now to the factor of velocity. Matters here are in one way simple, in another way neither simple nor well understood. There is little doubt in regard to the main facts. Both general observation and statistical evidence make it clear that when the quantity of money is expanding and prices are rising, velocity usually is greater than it is when things take the opposite course. There is in this a certain similarity to the tremendous increase of velocity which is occasioned by violent inflations. But the phenomenon of increasing velocity with moderately rising prices is not to be explained in the same way. In the extreme case of wild inflation, when people come to distrust money which they can still use and dispose of it as fast as they can,—then the element of confidence, or rather loss of confidence, explains the rise of velocity. But no such shattering of faith in deposit money occurs under the conditions of boom and optimism; and when a crisis does ensue, the circulation of deposits is not accelerated, but slowed down.

One cause of the acceleration in boom times is the factor of expectation,—the course which prices are *expected* to follow. When the movement is upward, business concerns, faced with the fact that the prices of raw materials and other supplies are rising against them, are likely to hasten purchases in order to get the best terms. They tend, therefore, to keep on hand less idle money, and to disburse what comes to them as soon as it is received. When the price movement is downward, on the other hand, there is the opposite kind of response. So far as the money used directly in busi-

ness is concerned, a rise of prices seems to promote its more rapid flow, a fall of prices its slower flow.

There are doubtless other causes. The whole matter of velocity and of changes in velocity presents a complex of forces which economists are only beginning to unravel. It would seem that the several elements do not in fact vary separately, but connectedly. Certain it is that we must be on guard against any literal interpretation of a quantity theory, and still more so against any offhand and mechanical resort to it either for the explanation of monetary experience or the determination of monetary policy.

§ 5. The importance of deposits in relation to the supply of money and the level of prices raises a further problem of great intricacy. Since deposits are created by banks when lending, and since the volume of the lending is affected by the rates of interest charged by them, there must be some connection between these rates of interest on the one hand and the supply of money and the price level on the other hand. What is the connection?

One might expect that the two would be inversely related. Low rates of interest, it would seem, would encourage a large volume of borrowing and hence lead to a plentiful money supply and a high range of prices; high rates of interest on the other hand lead to a small money supply and low prices. Curiously enough the facts seem to show just the reverse. Periods of higher interest rates are usually periods of rising prices, while times of low interest are times of falling prices. So much is a matter of statistical record about which there can be no disagreement. How is this apparent paradox to be resolved?

The question has been the center of much critical discussion among economists, but has not yet been answered in any way that commands general agreement. It reaches far afield into matters which have no immediate bearing upon the subject of the present chapter. It will be enough at this stage to consider the simpler theoretical aspects and then try to see how the general conclusions can be fitted with experience.

Other things equal, low rates of interest tend to swell the volume of loans and deposits and to raise prices. Business men finance their productive operations largely by borrowing from banks, and

they count the interest paid the banks as a cost to be balanced against the profit which production is expected to yield. Given, then, a certain expectation as to what the level of profits will be, the lower is the rate of interest charged by the banks for loans, the larger will be the net profit remaining to business men and the stronger the inducement to extend production by the aid of additional borrowed funds. On the other hand, if the rates charged by the banks are high—that is, high relatively to profits—the volume of borrowings tends to be smaller. Borrowing from the banks being so intimately connected with deposits, a rate of interest that is low relatively to profits tends to bring about an increase in the quantity of money and a rise in the level of prices, while a rate that is high causes a shrinkage of the quantity of money and a fall of prices.

The rates of interest charged by banks are in their turn affected by the movement of prices. Once the quantity of money is expanding and the range of prices moving upward, the fact that business men are then securing higher profits tends to cause interest to rise. In times when expectations are high and prospects for gain seem good, all business men want more "capital"; that is, all want the money means that will give them command of more capital goods and more labor. The demand for loans from the banks increases and the rates of interest charged by the banks are drawn upward. Conversely, when prices are falling, interest rates tend also to fall. The prospects of gain then seem poor, and the demand for accommodation from the banks is correspondingly slack. Thru it all, it is optimism or pessimism that chiefly counts.

This sort of connection between interest and prices appears to explain the correspondence between movements of interest and movements of prices. It is the expectations of gain which seem to account for the observed facts. Rising prices, a large volume of deposit money, high interest rates,—these go together; likewise do low interest, a small supply of deposit money and falling prices.

The wider causes bearing upon the situation are complex and obscure. Among them are the eventual accretions of profit and income, as contrasted with the expectations; the long-run rate of interest on investments as compared with the bank rates on loans;

the physical realities as compared with the money expectations. The relation which is most unmistakably seen in any statistical material is that just described.

It is to be remarked that in this proximate relation the connection seems to be closer between high rates of interest and rising prices than between low rates of interest and falling prices. Perhaps it would be better to say that the effects of *changes* in the rates of interest are more unmistakable in the first case than in the second. A familiar proposal toward mitigating the ups and downs of the business cycles, sometimes even put forward as a means of preventing them, is that bank rates of interest should be raised sharply in boom times in order to check dangerous activity, and on the other hand should be lowered radically in times of depression in order to re-induce and promote wholesome activity. A higher rate is expected to have a sobering effect when there is undue optimism, a lower one to be stimulating when there is undue pessimism. The two procedures, however, do not in fact seem to be on a par with each other. Apparently a sharp rise in a central bank's rate of discount, entailing a similar rise in bank interest all around, has more effect in checking a boom than the converse procedure has in promoting revival. This *seems* to be the case; the point is one of the many in the phenomena of the business cycle about which we have still much to learn. Probably it is another among the psychological phenomena which, tho hardly the main thing underlying the cycles, do play a large part in them, and at the same time a part which may easily be inconstant and unpredictable,—greater at one time than at another, greater in some countries and communities than in others.

Not least, it is to be remarked that this whole range of discussion refers to the bank rate of interest, and chiefly to the rate of discount,—the rate on the short-time loans in the day-to-day cold-blooded market as distinguished from the rate to a bank's regular customers. More will be said elsewhere on the various kinds of loans and the various rates of interest,—on long-period rates and changes in them, and on short-period rates and changes in these. It is the latter which are chiefly had in mind when we discuss the way in which the price level is connected with the rate of interest.

A different kind of analysis and interpretation must be applied to interest on capital as a social institution.[1]

§ 6. A direct relation between quantity of money and prices appears in the case of government paper money; a case of commanding importance in the twentieth century. Here the amount of the paper has in ordinary times a temporary and slender relation to the volume of transactions. That relation does appear in the fluctuations of the denominations of the notes, quite the same as the familiar fluctuations in the quantities of subsidiary coin.[2] But the total volume of government paper is quite independent of the course of business in the extraordinary times—war and the like emergencies—when resort to it is made on a great scale. As we have seen, the striking phenomena in the history of paper money appear in times of war inflation. Then there can be no question of the volume being the result of greater activity in business. It may cause greater activity of a wild kind, and this may lead to increasing issues; and then, as on other occasions of speculative outburst, there will be the demand for "more money in order to finance prosperity." No one would pretend, however, that the great historic doublings and quadruplings of prices in the course of wars have had as their fundamental cause anything other than the outpouring of paper money by governments.

This leads to a new element in the modern situation. It has been repeatedly remarked in the previous pages that bank notes are essentially similar to bank deposits, and that some of the characteristics and inconsistencies of legislation on banking in the nineteenth century and far into the twentieth have arisen because of the failure to perceive this similarity and the difficulty of adapting legislation to it when it has been perceived. But the similarity ceases when bank notes become virtually government notes. This they do become when the issue of notes is allowed only to the great central banks, as is now the case almost universally. True, in ordinary times there may be elasticity of issue, and adjustment of amount to the demands of the public for the various denominations of cash. There may also be ordinarily redemption of the

[1] See Chapters 38–40 in Vol. II.
[2] See Chapter 21, § 5.

notes in gold—maintenance of a real gold standard. But in times of stress all this is thrown to the winds. The notes become inconvertible paper. Their amount is a matter of the will of government and is not influenced by the volume of goods produced and offered. It rests on a quite extraneous factor, the immediate financial need of the public treasury. When this stage is reached, the premises on which the quantity theory rests are essentially in accord with the facts. If deposit banking also were all in the hands of government, the same would presumably hold of that form of money also. Then it could hardly be pretended that quantity of money and quantity of goods were interdependent; each would be an independent function.

§ 7. What now remains of the "quantity theory"? Not much, it would seem, as regards the matters which in the twentieth century bulk largest in the discussions of these matters both by economists and by the larger public. True, it is accepted as the main element in explaining the explosions from great issues of inconvertible money, familiar in history yet easily forgotten. It still has meaning, further, for problems which, tho forgotten quite as often, are of the highest importance: for the long-period movements of prices and so for the principles and the mechanism of a managed currency. It bears on the persistent problem of the desirability of retaining the gold standard. If that standard comes to be restored to something like its old place, the changes in the monetary supply of the metal will have to be taken into account, essentially on the lines of the quantity theory; a matter to be discussed further in the next following chapter. And if, on the other hand, the gold standard is discarded once for all, the only alternative is some sort of paper or token money. This will have to be watched and managed. There will be at least as great difficulty as under a gold basis in ascertaining in what way a moderate increase or decrease in quantity will affect prices. The same possibility will arise as of old that there may be violent changes whose consequences can be predicted only on the old lines. And if, further, the deposit system comes to be radically overhauled, any plan of action must rest on the assumption that less of deposit money means lower prices, and more of it higher prices.

When it comes to short-period changes in prices—not only the very shortest, but those for periods of some years—we are much in the dark about the extent and way in which the price level depends on the volume of money. Here the conclusions most tenable are largely of a negative kind. It is tolerably clear what will be in the end the cumulative effects of repeated additions to the quantity of legal tender money, whether gold or paper. But no one can predict with any certainty how soon the consequences will come or how great they will be.

Possibly the improvement of statistical science and of the materials on which the statisticians work will in time enable a first-rate government to know in advance just what measures will bring the monetary situation deemed ideal and how promptly they will do so. But this is no more than a possibility. The probability seems to be that for long-run plans we shall have to content ourselves with the old familiar doctrines,—truisms perhaps. It is striking that these truisms, notwithstanding the rough handling they have received in recent economic literature, slip inevitably into the discussion as basic when it comes to judgments and predictions at any given date on the complex problems of immediate monetary policy.

CHAPTER 30

THE GOLD STANDARD

§ 1. The essentials of the gold standard: the monetary unit convertible into a given weight of gold. Its older form: free coinage and free circulation.—§ 2. The later form: the gold bullion standard. Contrast to free circulation. Relation to paper money. A variant: the gold exchange standard.—§ 3. Two gains ascribed to the gold standard; stabilization of the foreign exchanges, and stabilization of the price level. The first clearly achieved; its importance sometimes exaggerated.—§ 4. Stabilization of prices more important but never so fully attained. How affected by the form of the gold system, by the sensitiveness of the banking system and by the character of international payments.—§ 5. Difficulties after the war of 1914–18. England. The United States. Other countries.—§ 6. The only alternative to the gold standard system is one of inconvertible paper. Both call for management, but the paper for more of it. Importance of psychology, of rooted traditions; of political factors and international relations. Some conclusions.

§ 1. THE subject of this chapter is closely connected with those of the chapters preceding, yet hardly less closely with the Book that follows. Obviously it is most intimately connected with money and banking; yet cannot be readily followed without some understanding of international trade. The reader may, if he wishes, postpone the present chapter till after reading the chapters on international trade. The two subjects, money and international trade, interlace. Yet a consideration of the history and meaning of the gold standard constitutes a sort of climax to what immediately precedes; and for that reason it is placed here.

The gold standard, its working, its advantages and disadvantages, its past and future, came to be looked at more critically after the great crisis of 1929 than at any previous time. During the nineteenth century gold was thought to be clearly the best basis for a money system; by the fourth decade of the twentieth century the prestige of the gold standard had been sadly impaired. Not only the quite discredited paper money visionaries of older days but sober and discriminating thinkers began to ask whether a better system could not be found, or at least some fundamental

changes made in the system itself that would remedy troubles supposed to be inherent in its traditional mode of operation.

The gold standard rests on the establishment by law of an equivalence between the monetary unit and a given amount of gold. Originally it was conceived to mean that the monetary unit should actually contain a stated amount of gold; for instance that the American dollar (by the law of 1834) should contain $23\frac{15}{21}$ grains of fine gold. In the course of the nineteenth century, as paper came to be used more and more, gold pieces were used less the world over, and in the twentieth century after the war of 1914–18 the metal served its monetary purpose only in the form of bars held in public reserves. Such phrases as "the weight of the gold dollar," "the gold content of the dollar" (or pound, or franc, or mark) are merely traditional, left over from former days. What the law means and what these phrases point to is that the circulating money is made exchangeable for a stated amount of gold. Under the laws of most countries the equivalence has been achieved in something like the old way, by naming a "mint price of gold." In Great Britain, for example, when the gold standard was restored (temporarily) in 1926, the mint price of standard gold (gold $11\frac{1}{12}$ fine) was fixed at the old figure £3 17s 10½d per ounce; but no gold coins were struck. Expressed in the same fashion, the mint price of gold in the United States, as fixed (for the time being) in 1934, would be $35 per ounce; but no gold coins were minted. It is a matter of indifference what mode of expression is adopted. As a basis for the *formal* establishment of a gold standard all that is required is that the law state the amount of gold in the monetary unit.

For the standard to be operative, however, something more is necessary. It is not enough to have mere legal enactment; the relation stated in the law must be established in fact. Indeed, no specific enactment is necessary, if there be adequate administrative power. The value of the currency unit must actually be kept equal to that of the specified amount of gold. This is the essence of the gold standard.

Under the older tradition, two things were thought necessary in order to assure this equivalence. First, the government must

stand willing to deal in gold in unlimited amounts at the legal (mint) price. Bullion or coin must be had on the stated terms by every person who presents paper or any legal tender other than gold, no matter how much he wants; and, on the other hand, gold must be bought at the same price in whatever amount it is offered. When gold coins are in use, this means "free coinage." Bullion must be convertible into coins at the mint, or coins into bullion, without limit as to amount. Where coins have ceased to be in circulation, and paper money takes its place, bullion must be given for the paper (or for checks payable in paper) without limit of amount. It is not necessary that the bullion be supplied without any charge. Where there is still actual coinage of gold, there may be a "seigniorage" to cover the expenses of minting. Where gold is not coined, there may be a handling charge or some such fee—corresponding to a seigniorage—to defray the expenses of conducting the transactions. An essential thing under the older conception is that the transactions be unrestricted.

The second principal requirement is that no obstacle be placed in the way of whatever disposal the holders of gold coin or gold bullion may choose to make of their metal. It must be permissible to melt down gold coin, if there be such in use; or, where only bullion is used, to obtain this from the authorities and to employ the metal for whatever purpose is desired,—e.g. for industrial uses such as the manufacture of jewelry and ornaments, the making of wire or the filling of teeth; these must be permitted without restriction. Most of all it is essential that there be freedom in the use of gold for export to other countries. One of the prime functions of the gold standard was to establish stable rates of exchange between the currencies of different nations, and this, it was believed, could only be done if there was no hindrance to the shipment of specie between one country and another.

§ 2. Contrast now the gold standard as it was thus in operation during the period of its heyday—the second half of the nineteenth century—with the form it came to have in the third and fourth decades of the twentieth century. The earlier form might be called the "classic" form, developed in the western world contemporaneously with the wide acceptance of the "classic" economic doc-

trines of Adam Smith and his followers. It is the system of free gold circulation. The later forms are the gold bullion system, and, as a minor variant of this, the gold exchange system.

In the classic form there was not only free coinage and unfettered access to coin and bullion at the mints, but also a substantial circulation of gold coin in everyday transactions. This, indeed, was often thought—half unconsciously, as if it were a matter of course—to be the only form in which a gold standard could exist at all. Gold, it was felt, had to be in hand-to-hand use.

To ensure that gold coin should be steadily in circulation, and especially to prevent it from being displaced by paper notes, various devices were adopted. One of the simplest of these, long practiced by European countries, was the prohibition of paper money—bank notes—of small denominations. Bank of England notes could not be issued under five pounds, Bank of France notes under fifty francs, German bank notes under fifty marks. In France and in Germany notes of even these larger denominations were in fact issued sparingly, with the express design of preventing the displacement of specie. Where such regulations were in force a considerable circulation of specie was assured. People need cash for everyday transactions, and if paper money is not available in the appropriate denominations, then gold coins will be called for and will remain in circulation.

The gold circulation system had disadvantages. From the standpoint of the money-using public it imposed some inconvenience. Gold coins in any but the smaller denominations are heavy and awkward to handle. On the other hand, when the transactions are of considerable size, notes are superior to gold. Again, from the standpoint of the money authorities, the system was expensive to maintain. The use of gold coins in everyday payments as well as for reserves against notes and deposits meant that a much larger amount of specie was required than otherwise would have been the case. Not only immediate expense but another ground, of wider economic weight, prompted a more sparing use of the metal,—fear of a coming shortage in the world's supply. These considerations, together with others of a more dubious sort, such as the supposed advantage for political and military power of a

great centralized stock of gold, put the gold circulation system in disfavor. After the war of 1914–18, when people thruout the world had become well accustomed to the use of notes of all denominations, a different form of gold standard came into use.

Under that newer form, specie is used in the form of bullion only. Instead of being made up into coins, it is kept in bars; not put into active circulation but confined in reserves. It is dealt in freely by the monetary authorities at the legally established price, but it is not minted. Under this, the gold bullion system, the metal still acts as the standard; the value of money is kept equal to that of gold. The overt change that has taken place is only in the shape in which gold appears.

The gold bullion system involves a clear recognition of the fact that where notes of a satisfactory sort are issued the circulation of specie is unnecessary. As long as bullion is readily available for industrial uses or for shipment to foreign countries, the full gold standard can be maintained. There is no need to mint coins. The purposes for which gold is wanted are just as well served by specie in bars.

The confinement of gold to holdings in monetary and banking reserves is the latest stage in a long process of evolution. Originally used as the medium of exchange because its sheen and its durability satisfied the universal desire for display, it was gradually displaced in the familiar hand-to-hand uses. First came bank notes; then considerable bank deposits; and finally (and it would seem definitively) paper money issued or controlled by government. It is something of a paradox that a metal first used as money because of its appeal to the eye should continue to be so used as money only in the form of unsightly bars buried in the vaults of central banks and treasuries, or in that huge anomaly, Fort Knox.

A third form which the gold standard has taken, an attenuated variant of the bullion system, is the gold exchange standard. It requires, curiously enough, no holding of gold whatever on the part of the country that maintains it. In lieu of reserves of the gold itself, reserves are kept in terms of some important foreign money that is fully established on a gold standard. The reserves may be in the form of deposits held in the banks of the foreign country,

or in bills of exchange or other claims readily convertible into the foreign money. So long as the foreign money itself is in fact exchangeable for gold, the country operating the gold exchange system can keep the value of its currency equal to that of a fixed amount of gold, simply by buying and selling the foreign money (or claims to it) at a fixed rate. In other words, it can maintain a gold standard for international payments and for the foreign exchanges. Thus for many years before the war of 1914–18, the Indian government kept the value of the rupee at about 1s 4d with respect to sterling, and after the war kept it at about 1s 6d. As long as the pound sterling was convertible into gold at a fixed parity, the rupee was too, because convertible into sterling exchange.

The great advantage of the gold exchange system is that of economy. From the standpoint of the public treasury and the public finances, the deposits in foreign banks and the bills of exchange on foreign countries, besides fulfilling the function of reserves, stand also as a source of gain; they earn interest. For these reasons it might be thought that the gold exchange system would be in wide use. But it has not been. So strong is the popular feeling in most peoples against anything which appears to signify a dependence on other nations that only a very few countries, and these not of the first economic importance, turned to the gold exchange system with an expectation of its permanent retention. Its main use has been as a first step toward the restoration of a gold standard after periods of monetary disruption. It was so utilized by Austria in the nineteenth century and by other European countries after the currency chaos of the war of 1914–18; and it may be used again in this way.

§ 3. In considering both the past and the future of the gold standard, a distinction should be drawn between two benefits which it may bring. The benefits are: first, the stabilization of the foreign exchanges; second, much the more important of the two but also the less easily attained, the stabilization of the price level. The first end is reached in much the same way and much to the same extent, by the gold standard in whatever form; while there are important differences as regards the second.

Where the moneys of two or more countries are established on a

gold standard, movements in the foreign exchanges are held within very narrow limits. This subject will be dealt with at a later stage,—the par of exchange, the conditions under which gold flows from country to country, the narrow limits within which the rates of exchange fluctuate. There is no substantial deviation from accuracy if we say that where the trading countries have the gold standard, their rates of foreign exchange are fixed.[1]

Fixed exchanges immensely facilitate all international economic intercourse. They render possible a greater substantive movement than would otherwise take place. As regards the great staple commodities which make up the bulk of the tonnage movement between countries—ores, grains and sugar, cotton and wool and jute—trade operates on an extremely narrow margin and large transactions are undertaken which promise only a small percentage of gain. So long as the dealings can be conducted with the assurance of even a tiny gain per unit they are carried on freely. Fixed exchanges enable dealers and merchants to make close and accurate comparisons of prices in domestic and foreign markets and to dismiss from their estimates of profit the chances of great ups and downs in the foreign exchanges. Without stable exchanges, movements of the exchange rate are unpredictable, and merchants have no way of calculating with certainty what dealings will be worth their while. As a result, the volume of transactions is reduced.

How great is this effect—how much the physical volume of international trade is lessened by erratic foreign exchanges—is necessarily a matter of judgment and guess. Very great and rapid fluctuations no doubt have a long-continued injurious effect, may even have a paralyzing one. But it is not clear how much damage is done by fluctuations which, tho rapid, are kept within a moderate range. Brokers and bankers specialize in the foreign exchanges, and thru them the merchant can cover himself, buying his exchange once for all and leaving the middlemen to do what they can to minimize their own risks and maximize their gains. The same sort of thing happens in domestic trade where there are organized exchanges and large speculative transactions in com-

[1] See Chapter 31.

modity markets; and the same sort of doubt may arise whether an elaborate marketing organization is not attended by unwelcome consequences. In both cases—in domestic as well as international speculative operations—it would seem that the physical volume of transactions remains ordinarily much the same.

The situation is different where the irregularities and disturbances of the exchanges are not only great but persistent. In the course of the years of depression that followed the crisis of 1929, the conditions of international trade were quite anomalous. The physical volume of the trade was reduced to a surprising and disrupting extent. And the great dislocations of the foreign exchanges were in themselves a cause of the physical decline; for they gave occasion for higher import duties, restrictions on imports and exports, quotas, clearing arrangements and the like. These hindrances were not the consequence solely of the exchange fluctuations; other factors entered, such as the general fall in prices, political dangers, panic flights of capital first one way then back again. It is impossible to separate effect from cause; to decide how far rapid and uncertain changes in the rates of foreign exchange were causes in these anomalous happenings, how far results. All that can be said in general terms is that there is some substantive gain from the simplicity in the foreign exchanges which comes when all countries maintain the gold standard, and comes even when all the more important countries do so and others do not, and that there is some substantive loss from repeated spasmodic fluctuations.

Much the same is to be said of another substantive aspect,—the movement of capital from the older countries into the newer. That takes place more smoothly under a common standard, and the gains which accrue to the lending and the borrowing country are secured at less cost and are subject to less disturbance. Where there are dislocated foreign exchanges, the disturbing element of risk enters no less into capital transactions than in the import and export trade. Indeed, it enters more, since the capital contracts are in general for long periods. Commonly it is sought to eliminate the risk for the lenders by specifying payment of interest and principal in terms of a given weight of gold—the weight tradi-

tional in the coinage systems of one or the other of the countries. But stipulations of this kind have proved difficult, indeed virtually impossible, to enforce in the courts. For capital movements even more than for merchandise transactions great and rapid fluctuations in the exchanges, brought about by far-reaching political and social disturbances, have serious effects on the substantive situation.

§ 4. It is stability in prices that is much the more important thing. This was supposed to be greatly promoted—even attained so far as was humanly possible—by an international gold standard. How far was the result in fact attained? And how completely and how smoothly? The answer to these questions is not the same for the earlier and the later forms of the gold standard,—the free circulation system and the bullion reserve system.

The expectation of stability in prices was based, as need hardly be said again, primarily on the slowness of changes in the total supply of monetary gold; and further on an expectation that the total volume of transactions also changed but slowly. This latter expectation was not often specifically referred to; it was held in mind half-unconsciously, being a common-sense generalization not easily verifiable by statistical evidence, but tenable, and on a little reflection even obvious. The first element, the slowness of changes in the monetary supply of gold, was deliberately and elaborately considered, and played a large part in the literature of the subject; and it has been sufficiently dealt with in the preceding chapters.[1] What concerns us here is that the slowness of change in gold supply was supposed to exist not only in the world at large but in each country which had the gold standard. It must be borne in mind that the thing supposed to be constant was the monetary stock; not the total stock, still less the annual product from mines.

Now the stability of gold supply, and the promotion thereby of stability in the price level thruout the field where the gold standard obtained, rested on the further supposition that gold moved freely from country to country, and that all countries thus were one as regards the general movements of prices. Any unusual rise of prices in a single country served to attract goods from other

[1] See Chapter 21.

countries and so checked a rise, while any fall in prices in a single country would attract gold and so check a fall. The effects of any change in the monetary supply of the metal would be spread over the entire system. Even tho a change in volume were such as might bring sudden and great price movements in a single country if confined to that alone, no violent disturbance was to be feared if the effects were spread over the whole field.

All this seemed to be proved by the experiences under the great historic changes in gold supply, especially by the outcome of the Californian and Australian gold discoveries of the middle of the nineteenth century and by that of the South African discoveries at the close of the century. In both cases the world level of prices rose, and the movement was slow for the world at large. True, the rise was made slow largely because of the adventitious circumstance that the gold was not a *net* addition to the monetary supply; in good part it served to displace silver which flowed to the Orient. A marked increment nevertheless there was in the Western world, and the tendency of prices was to rise. And yet, notwithstanding changes in gold output on a scale never before known, prices rose at a pace very moderate. And the rise, so long as it was moderate and gradual, might be deemed advantageous as giving a stimulus and a zip to industry. These two periods of rapidly increasing gold supply, both coming in the full flower of the gold standard, were periods of marked prosperity and marked increase of international trade in the world at large. Not only did prices rise but also the physical volume of production—the national dividends in terms of goods and services.

This smoothness of the movement over a great area was promoted by the ready impact of the gold flow on prices in the several countries. Evidently the more sensitive the kind of gold standard possessed by a given country, the more certain and rapid is that country's response to an inflow or outflow of the metal. If all countries are sensitive, the more likely is it that the price level in the world at large will not change rapidly. This very sensitiveness, tho a factor that makes the gold standard work better, is commonly felt to be a disadvantage where it leads to a check on rising prices, still more so if it happens to work in the direction of a

positive fall in prices. Hence the endeavor to accumulate in a central bank a large reserve of gold, which should be available for remittances abroad without bringing immediate pressure to bear on the commercial banks, or—if pressure cannot be averted—with all possible easing of the strain.

Of prime importance in the matter of sensitiveness, and thru this of the spreading over a large area of the effects of changes in gold supply, was the position which gold so long held as everyday money. What this means and what follows from it has already been pointed out. Where the circulating medium is made up largely of gold coin, the sensitiveness is greater than under the bullion system; for an outflow or inflow of the metal impinges more quickly on the monetary structure. True, where there is a wide use of deposits, the effect is in neither case at all instantaneous. Gold that flows into a country then goes first into bank reserves; gold that flows out comes from them. While the changes may affect deposits, the rapidity and magnitude of the effect depend on the abundance of the specie holdings, on the extent of their concentration in central banks, and not least on the state of industry and trade at the time. But if deposits do increase and prices do rise as reserves are swelled by gold inflow, the commercial banks are also called on to supply more hand-to-hand money to their customers. Gold being in hand-to-hand use, a part at least of any new supply filters into circulation. The converse ensues when there is a falling-off in deposits; gold piles up in the banks. In general, a gold standard with free circulation and readily available bank reserves has a certain automatic balance; the reserves are subject to internal drain when imports of specie push prices up, and accumulate when prices fall and so serve as the basis of the next upward movement. The bullion system is less sensitive than the gold circulation system, and for that reason is more smooth; but it also is less automatic in its working and calls for more constant and more difficult management.

Sensitiveness, tho one of the factors which made for the world stability of prices under the gold standard, was often felt to be a source of danger; or if not dangerous, at all events inconvenient. In times of boom it was said that there was "not enough money

to finance prosperity"; which meant that more cash was called for in everyday circulation and that the banks did not relish the pressure to provide more and more of it. Hence too the familiar demand for a "flexible currency"; heard little in times of quiescence but loud when the business stream was rising dangerously. It is an old and familiar story. People approve of salutary regulations in the abstract, but wince when the restrictions come to be applied to their own doings.

There are other grounds, more valid, for steps toward lessening sensitiveness. It renders the whole monetary system, and the whole industrial system, vulnerable to strains which may be recurrent and at the same time unpredictable. They arise more particularly from the extraordinary growth of international payments other than for merchandise imports and exports,—the "financial" operations; and also from purely political causes. Sudden strains on both accounts were not uncommon before the war of 1914–18. After it, they became much more common. Sales of the well-known corporate stocks and bonds and of public obligations could be made at a moment's notice in any part of the world. Transfers of them, tho sometimes they served to take the place of specie movements, were also causes of such movements on a great scale. These factors in international payments, already important before the war of 1914–18, not only became of larger consequence after the war but were added to and complicated by the great mass of political debts and payments between countries,—the interallied debts and the reparations. Further complications arose from the political uncertainties of the post-war period itself. The fear of monetary changes, and also of instabilities, nay revolutions, in governments, led to panic sales and transfers, sometimes into gold-standard countries, sometimes out of them when devaluation was feared.

Where international payments arise chiefly from the import and export of goods, there is little likelihood of sudden impacts on a gold reserve, or of unexpected disturbances. Disturbances there may be, but none so sudden and sharp as those likely to arise from the financial and security operations, or from the rumors and panics of political overturns. As matters stood in the third and fourth decades of the present century great international

remittances were called for, sharp fluctuations in the foreign exchanges arose, gold moved to and fro in great amounts; all of which had no immediate connection with the imports and exports of goods, with real business.

Conditions of this kind led inevitably to endeavors to make the monetary systems *less* sensitive; to protect the gold standard and its basis in central bank reserves from great shocks that were not only irregular but often temporary. The establishment of the Federal Reserve system of the United States in 1913 was largely brought about by this wish. Gold reserves then were concentrated, and were made so large (unexpectedly large, as it proved) that an inflow and outflow could take place without any impingement on the internal monetary structure. In England, the same sort of policy, already advocated and started in pre-war days, was explicitly adopted when the return to the gold standard was undertaken in 1926. Then England gave up for good the sensitive gold circulation system and turned to the less sensitive gold bullion system. France did the same when returning to the gold standard with the devalued franc of 1927. The effect of the somewhat tortuous steps taken in the United States in 1933–35 was the same; gold bullion was piled up in huge amounts in government vaults, and no gold was allowed to circulate either in the form of coin or of gold certificates. Slowly at first, but with a rush during the first decade of the post-war period the bullion system became the only form of the gold standard that was in effect.

What all this might lead to was not fully realized as it was going on. The free movement of gold from country to country was impeded, directly or indirectly, and when it did take place the effect on prices and trade was uncertain and slow. The relations between the price structures of different countries were no longer close. The influence of the gold standard toward smoothing the distribution of the monetary gold between different countries and thereby conducing to stability both in the international movements of prices and in the domestic movements became slow and uncertain.

§ 5. The lessened sensitiveness of adjustments thru gold flows

and the increasing volume and spasmodic character of the financial payments made the working of the gold standard more and more unsatisfactory. Some illustrations of these difficulties will make more clear the new situation which emerged after 1929.

Most instructive is the experience of Great Britain. In 1925 the gold standard was reëstablished in the bullion form. It was the first explicit resort to that form, soon copied in the other countries that returned one after one to a specie basis,—France, Italy, Germany, Austria. Yet in 1931, after but five years of experience, it was given up. The causes of this misadventure need not be considered in any detail. The main thing was that, when the gold standard was re-established in 1925, the adjustment of domestic prices to the new conditions proved so difficult as to be thought impossible. The old parity had been maintained and it proved too high. Success in the maintenance of a restored gold standard depends during the first years after re-establishment on a favorable, or at least not unfavorable, balance of international payments; and this depends in turn on a maintenance of exports relatively to imports. Over a longer period, if the price level has not already fallen to the range tenable under the gold standard, there must be a lowering of the domestic income level and of money wages,—"costs." But the laboring classes in England resisted stubbornly any reductions in money wages, as people of all sorts and kinds invariably do. The strength of the Trade Unions, the circumstance that the Labor Party was in power, the established practice of government participation in negotiations and agreements about wages,—all these made it politically impossible to insist on lower wages, or to let them come (as was the way in former days) by a long and wearisome process of strikes, lock-outs, compromises, and in the end acceptance of the inevitable. This "stickiness of costs" (a term which means in the main a clinging to an existing rate of money wages) was by no means a new factor, but proved to be of more weight and to have wider consequence than ever before. More than anything else it led to the monetary collapse of 1931.

It is true that other causes of a more adventitious sort increased the difficulties. Panic conditions in 1930–31 on the Continent led to heavy withdrawals of funds from England, the one country

whose tradition it had been to pay out gold whenever called for. Probably the suspension of gold payments would have come sooner or later even without the difficulties which thus happened to beset the Bank of England in 1930 and 1931; they were hardly more than the immediate occasion of the step finally taken in 1931. Great Britain, in thus going back on her tracks, paid dearly for the manful obstinacy which had led her to resume specie payments in 1925 at the former mint price of gold and to face courageously the possibility of a depressing downward trend of domestic prices and of money wages.

In one important respect the difficulties of the new move were smoothed. Tho it led to dislocated foreign exchanges, some degree of stability in the exchanges was secured thru large government purchases and sales of bills; not indeed stability like that arising from a fixed mint price of gold but a prevention of abrupt fluctuations and an adjustment of the general trends according to the real or supposed exigencies for the time being. Internal prices were watched rather than regulated; at all events they were not longer left subject to a steady downward pressure.

Very different was the course of events in the United States. Here during the third decade the gold flow on the whole was into the country, not out of it; and the gold standard was at no time in danger. The inward gold flow was the outcome of various influences. A strong demand for American goods abroad, coupled with an intensified policy of high protective tariffs against imports, led to a large excess of exports and tended to cause gold to flow in. There was an increased flow of interest payments on capital which had been lent abroad, and for a while there was a flow of repayment of war debts owed to the United States government. To some extent the tendency was offset by a large volume of new American investments abroad; indeed, in some years the new investment was large enough to turn the balance the other way. But on the whole the trend was toward gold imports into the United States. The monetary gold stock of the country, already large at the end of the war, rose to enormous amounts.

Yet the officials of the Federal Reserve system and those of the government hesitated to allow any rapid expansion of the cur-

rency system. Any great expansion, it was felt, would lead to the familiar round,—a financial crisis followed by a depression. That this outcome actually came, notwithstanding the hesitating efforts to prevent it, serves perhaps to justify the general policy of the authorities. At any rate, the view held by them, whether wise or unwise, led to an attempt to put a curb on monetary expansion. But the collapse of 1929 showed no resemblance to the traditional working of the gold standard. The speculative mania which preceded the collapse, so far from having been held in curb by an outflow of the metal, served rather to promote its inflow; while the huge accumulated fund of the metal, and the powers in the hands of the Federal Reserve system, great as they seemed at that time, were of no avail for preventing the crash of 1929 or the breakdown of the entire banking system in 1933.

The final outcome in the United States was thus quite out of accord with traditional expectations, or indeed with any reasonable expectations. Tho the holdings of gold in the Reserve Banks were enormous, and quite beyond what was thought to be a safe strong reserve; indeed it was widely thought that the stock was so far in excess of any domestic needs that much of it might be shifted somehow or other to foreign countries which had difficulty in clinging to the gold standard—payments in that metal were suddenly discontinued in 1933. The gold standard at the old mint price, which had been maintained for over fifty years since the resumption in 1879 under extraordinarily varying conditions of gold scarcity and gold abundance, was suddenly cast aside. What was virtually a régime of inconvertible paper followed. Foreign exchange, no longer kept stable within the narrow limits of the old gold points, was regulated much in the English way, by government purchases and sales of gold, at prices (virtually in terms of paper money) which were fixed from time to time at the discretion of the Treasury acting thru the Federal Reserve Bank of New York as its agent. In 1934–35, as has been noted, something like a firm mint price of gold was established, and so something like a return to the gold standard. This was a sort of bullion standard; but it rested not on a specific law of Congress but on administrative action subject to change at any time with merely an inti-

mation that a deviation from it was not likely to come at any early date. The grounds for the original drastic move, that of 1933, have already been considered,—a wish to make prices move upward, and quite erroneous notions as to the extent and rapidity of the effects on prices to be expected from a lessening of the amount of metal in the gold dollar.

Whatever the causes of the several steps taken in the United States between 1933 and 1935; however they were different from those taken in England in 1931 and thereafter,—the outcome in the two countries was much the same. The combined effect was to undermine the old traditions of the gold standard. As regards free circulation, that standard would seem to have been irretrievably shattered. How far it was destroyed once for all—even in anything like the bullion form—no one could tell.

In 1936 a number of countries which had adhered to the gold standard succumbed to the prevailing currents. Some of them—France, Belgium, Italy—had already "devalued" in the preceding decade, when returning to the gold standard with a new and smaller "gold content," i.e. a new and higher mint price for their money unit. France and Belgium now devalued once more; and Italy, which had established a sort of *de facto* gold standard, modified this by a second devaluation of the same extent as that of the other countries. The only countries which had never suspended or devalued at all, like Switzerland and Holland, felt compelled to join the prevailing rush and proceeded to devalue. There was a widespread belief that a country loses something precious from the mere parting with gold; and a further belief, equally widespread, that devaluation has a special effect in enabling exports to be sold more readily than before and in preventing imports from coming in as readily. Both the extent to which shifts of this kind are likely to occur and their duration are apt to be exaggerated. The effects are greater and perhaps more lasting in a country whose international trade bears a large proportion to the total of its industry and exchanges (as Belgium) than in one where the proportion is small (as in France and the United States). In any case the abiding forces that determine the course of substantive trade between nations are not modified; tho it is

quite possible that they may be lamed by the repeated intervention of the various devices for regulating that trade which were tried during the great depression of the thirties. The immediate exigencies and the immediate expectations affect legislation in this field more than in any other, and often leave an unwelcome aftermath when the fears and hopes of the time have been quite forgotten.

§ 6. The only thing that can take the place of the metallic standard is some system of inconvertible paper money. And such money is necessarily dependent for its good working on management once for all.

No doubt gold money, tho it has a certain automatic character, has always been in an important sense a managed form of money. Coinage may be described as merely a kind of management. Under the gold bullion system the working becomes less automatic and continued management becomes more important and more necessary. And yet there is a kind of steadiness in a gold basis for money which it is hard to secure under any other system, and which makes it less dependent on government support than paper money can be. Sober and conservative traditions are attached to it. Notwithstanding the suspensions and resumptions, spasmodic devaluations, gold continues to move from country to country, and not least between those which have given up the gold standard or have greatly weakened or attenuated it. There remains a widespread and deep-rooted feeling that money should be equivalent to, or convertible into, something durable and substantial. It should be more than a piece of printed paper. Traditions of this kind are especially valuable in new countries which are trying to advance fast, to build up their resources, to secure ample funds for investment; countries in which there is a constant temptation to make more "capital" available by printing more money. It has long been one of the commonplaces of economics that the danger of overissue constitutes the essential argument against inconvertible paper, while it is the comparative stability of the supply which underlies the argument for the metal gold. Obviously, the continued use of the metal as the monetary basis is a means, not an end; the end is the stability of the price level. But here, as with so many human

institutions, the means are more likely to bring the end if they are backed by a well-tested tradition,—"solid gold."

Nevertheless, the question of the maintenance or reëstablishment of the gold standard is, to repeat again, one of management. Will a pure paper standard or one resting on gold be managed better in the long run? The answer will be different for different countries and for different times, and cannot be a simple or unqualified one for any country or any time. For Great Britain and the British Dominions the long-standing sober and steady traditions both of currency legislation and of public administration make the outlook for good management better than in any other great economic and commercial unit. In the United States the chances of thoroly bad management seem to me less now (in this the fourth decade of the twentieth century) than they were fifty years ago; even a prolonged war would hardly lead to wild inflation. The two English-speaking communities might arrange between them for stabilized foreign exchanges, avoidance of spasmodic changes, ease and simplicity of international payments over their combined areas, a smoother course of industry and trade within each country and between them all. Further, if they did so, other countries might gradually join; and even those remaining aloof and going their own way would secure at least part of the benefit.

Obviously these are matters not solely of economic reasoning and experience. Political relations between countries and political situations within countries are of enormous importance. Even when we remain within the familiar range of economic theory and economic policy it is difficult enough to reach conclusions with certainty. We are not on sure ground as regards the merits of various kinds of stability in prices and incomes. The best informed and most experienced cannot be sure what kind of political and administrative set-up is most serviceable in monetary and banking matters within a state; still less sure what is the best way of moving toward international coöperation. An analysis wider than that traditional in economics would have to be undertaken: one that included national feelings and national rivalries, political institutions and traditions, and the psychology of the case. Not least,

the possibilities of war must be faced. War in the twentieth century means the disruption of all economic life, and most surely and quickly that of monetary systems. When war breaks out, the question is not what sort of monetary system is the best but what sort is the least bad. Comparative badness again is then a matter not only of minimizing the evils of the state of war itself but of also keeping in mind an eventual return to sanity. Sanity in monetary legislation usually has meant and apparently still means first of all a return to the gold standard. It is from this point of view that the strongest case can be made out for that standard. It has a kind of psychological as well as material sturdiness which paper money cannot possess. The quantity and quality of gold are felt by all to be not merely a matter of the printing press.

No one can foresee now what the world will look like in 1950, still less in 2000. It may settle down into a state not essentially different (even tho much modified) from that of the opening of the twentieth century; or it may have been changed beyond recognition. If the existing economic order goes, so will it be with its monetary systems. If an era of peace and good will emerges after the welter of 1930–40; if the foundations of private property remain, with merely a new and better superstructure—then the monetary system too may again be like that of old: with a metallic base so widely accepted and supported as to be firm and with stability in prices and incomes as great as can be expected under any system. Management there must be in any case; and better methods can hardly be secured otherwise than by trial and error. Given the same care and understanding, an international gold standard is as likely to be well managed as a heterogeneous set of purely domestic paper monies.

REFERENCES ON BOOK III

Large as is the literature on money and monetary phenomena, there is no first-rate book on the subject by itself; which indeed is almost inevitable, in view of the close connection between this and almost every other branch of economic inquiry. Outstanding among the chapters on the subject in general treatises are those constituting Book I of Marshall, *Money Credit and Commerce* (1922). A keen and stimulating volume is R. G. Hawtrey, *Currency and Credit* (1923); another, no less

so, is D. H. Robertson, *Banking Policy and the Price Level.* The two great books of J. M. Keynes, *A Treatise on Money* (1930), and *The General Theory of Employment, Interest and Money* (1936), reach far beyond the subjects traditionally discussed under money. Thoroly original and of the first intellectual quality, written by a master of style, yet not always easy to follow, they have stirred a great volume of discussion, criticism and explanation. It is not now (1939) to be foreseen how much they will eventually reshape economic theory.

On banking W. Bagehot, *Lombard Street* (1873), is still to be read, notwithstanding the changes in monetary conditions since its date. C. F. Dunbar, *The Theory and History of Banking* (first edition, 1891, 5th edition revised and brought to date by O. M. W. Sprague, 1929) is compact and good thruout. W. R. Burgess, *The Federal Reserve System and the Money Market,* gives an excellent description and analysis. C. A. Phillips, *Bank Credit* (1920) is valuable for its analysis of the deposit structure.

On government money, the remarkable German experience is described and analyzed in F. D. Graham, *Hyper-inflation in Germany 1920–23* (1930). On the gold standard, a critical and objective survey is in A. D. Gayer, *Monetary Policy and Stabilization* (1937).

The literature on index numbers of prices is large. Noteworthy among earlier contributions are those of W. S. Jevons, gathered in his *Investigations in Currency & Finance* (1889); F. Y. Edgeworth's Reports of 1887–89 to the British Association for the Advancement of Science, reprinted in his collected *Papers Relating to Political Economy,* Vol. I (1925). An able and exhaustive analysis of the various kinds of index numbers and of the grounds for choosing among them is in I. Fisher, *The Making of Index Numbers* (second edition, 1924). With this compare W. M. Persons, *The Construction of Index Numbers,* 1928. A summary showing the statistical methods and the results for various countries and various classes of goods is by W. C. Mitchell, *Index Numbers of Wholesale Prices in the United States and Foreign Countries,* Bulletin Number 284 (1921; reprinted as Bulletin 656 in 1938) of the U. S. Bureau of Labor Statistics. An excellent brief discussion is in E. E. Day, *Statistical Analysis* (1925).

BOOK IV

INTERNATIONAL TRADE

CHAPTER 31

THE FOREIGN EXCHANGES

§ 1. The "foreign exchanges," based on the varying coinage systems of different countries. In this chapter considered on the supposition that the gold standard prevails thruout. How bills of exchange settle payments without the movement of specie.—§ 2. The par of exchange, and premium and discount of exchange; illustrated by sterling exchange in New York.—§ 3. Bankers as middlemen in the foreign exchanges. Fluctuations in rates, due to the higgling of the market.—§ 4. Dealings between a series of countries, illustrated by transactions between the United States, England, and Brazil. The wide use of sterling bills for trade between all parts of the world.—§ 5. In what manner prices are influenced: in the long run, by the flow of specie; for shorter periods, by the rates of discount. Various complicating factors.

§ 1. The mechanism of international trade is not essentially different from that of domestic trade. It is part of the ordinary machinery of exchange; and it is closely connected with the banking operations and monetary phenomena of the several countries. The theory of international trade presents no fundamental peculiarities: it is but a phase of the general theory of exchange value. But it has been so much debated, is so beset by political and national prejudice, and is so peculiarly tinged by error in popular discussion, that there is advantage from treating it separately.

International trade, like virtually all the trade of modern countries, is carried on in terms of money and thru sales for money by individuals. Like all other trade, it brings in the end the same result as barter—the exchange of goods or services for other goods or services. But proximately it means sales for money. We may advantageously begin our consideration of it by taking up first the money mechanism thru which it is carried out.

That money mechanism may be of two kinds. The trading countries may have monetary systems resting on the same base, or they may have monetary systems quite unrelated to each other. The first case is that of the gold standard as it stood in Western Europe before the war of 1914–18. The second case is that of dislo-

cated currencies; chiefly the varying kinds and forms of inconvertible paper money which came into use so widely after the war. For simplicity and convenience it is best to disregard for a while the second case and to confine attention to the first. The essential characteristics of international trade are the same whatever the monetary mechanism; they merely stand out with greater clearness where the monetary systems have the homogeneity which comes when the gold standard is widely used. In the present chapter of this Book and in those immediately following, we shall examine what happens under the comparatively simple conditions, and thereby dispose of the fundamental questions. A later chapter of this Book (Chapter 33) will then take up the questions which arise under dislocated exchanges,—by no means unimportant, yet not such as affect the validity of the conclusions previously reached on fundamental questions.

When a merchant sells goods to a person in the same country the mode of payment is simple: he receives the money of his own country. But when he sells to a person in another country, it is not so simple. Transactions in England are settled in pounds, shillings, and pence; those in the United States in dollars and cents. The American who sells in England may sell there in terms of English money; he must then convert the English pounds into American dollars before they are available for him. Or, if he sells in England in terms of American money, he puts the English purchaser under the obligation of converting into dollars the pounds which alone are current in England.

This process of converting the money of one country into its equivalent in the money of other countries is carried out thru foreign bills of exchange. Strictly speaking, a bill of exchange is simply an order by one person, addressed to another, directing a payment to be made to a third person. It thus has three parties: the maker or drawer, drawee or acceptor, and the payee. When made out in the precise form settled by the law, it establishes a guarantee by the maker to pay the stated sum, in case the drawee does not do so; and when accepted by the drawee (he thus becoming the "acceptor") it fixes unconditional obligation upon him to pay it when due. Bills of exchange were at one time freely used in

domestic transactions, and were known as inland bills. A check is but a kind of inland bill of exchange, drawn by a depositor on a bank in favor of a third person. Foreign bills of exchange have no legal peculiarities. Their economic peculiarities arise only from the differences in the currency systems of the various countries. In the following pages, when bills of exchange are spoken of, foreign bills will be meant. The mechanism of payment in foreign trade thru such bills is usually called "the foreign exchanges." That term might as appropriately be used to describe exchange between different countries in all its phases, but is limited by custom to the dealings in foreign bills, while "international trade" commonly refers to the substantive thing,—the flow of goods from country to country.

For simplicity in exposition, let us begin by assuming that the only transactions leading to the use of bills of exchange are those by which goods are sold. There are other transactions of much importance, but the main principles are most easily explained in connection with merchandise transactions, and it is these which are by far the largest in volume.

Let it be assumed also that the gold standard prevails in the trading countries, gold being freely coined in them and flowing with ease between them. The complications which arise where there is not the gold standard must be reserved for later consideration; they have come to be of much importance, but do not lead to modification of the essential principles of international trade. It is curious that gold, tho deprived of its former position of dominance in the domestic transactions of practically all countries, still is accepted universally in settling international obligations and moves from one country to another in much the same way as before its dethronement.

A merchant in New York who sells goods to a merchant in London has a claim to receive money from the latter; he can draw on the Englishman for the price. He can draw directly or he can transfer his right. Thus it is the exporter who has bills of exchange for sale. On the other hand, a merchant in New York who has bought goods from a merchant in London has an obligation to pay money to this Englishman; he must remit in some way the

price. That is, an importer needs to buy bills of exchange. We are supposing here, again for simplicity, that both transactions are carried on in New York; the exporter sells his bill on London in New York, the importer buys his bill on London in New York. Suppose now that the two obligations are for the same amount, say £1000. The importer can buy from the exporter the latter's bill, drawn on his London debtor for that amount. The importer sends the bill to his London creditor; the latter collects the sum from the London debtor. The New York creditor gets his money from the New York debtor, and the London creditor gets his money from the London debtor. By one payment in New York and another in London, the transactions are liquidated without any sending of specie from one country to the other. Thru the mechanism of the bill of exchange, the exports serve to pay for the imports.

§ 2. What amount, now, would the New York importer pay, in American money, to the New York exporter? One thousand British sovereigns contain as much pure gold as $4866. Hence when a bill for £1000 sells for $4866 or its precise specie equivalent, exchange is said to be at par. If the American creditor sent to England for his money, brought the gold from London to the United States, and had it coined into American dollars, he would get from the mint this exact number of dollars, $4866.

Suppose, now, a number of exporters and importers in both countries, and a large volume of dealings: the case remains the same. The exporters sell bills, the importers buy them. If the money value of the imports just equals the money value of the exports, the bills of exchange exactly liquidate the transactions. Under such circumstances, exchange will be at par. Foreign trade will be in a state of equilibrium, the exports will just pay for the imports, and no specie will flow from one country to the other.

Suppose, next, that for some reason the exports from the United States exceed the imports in money value. The two sets of transactions—the buying of goods from persons in England, and the selling of goods to persons in that country—are quite independent. The American exporters may sell goods to a greater money value than that of the goods which the importers have to pay for. They

will then offer bills for a greater amount than the importers have occasion to buy. Under these circumstances all the bills cannot be sold to importers; some will necessarily be left over. The exporters who have the excess on their hands can do nothing but send to England for the specie. This, however, involves expense. The specie must be checked with care, must be boxed, insured, transported by land and water. When it reaches the American creditor, it must be carried to the mint and coined into American dollars—a process which may take some time. There is the possibility that some of the sovereigns may not be quite full weight, even tho not below the limit of tolerance in England. Not least, there is a loss of interest during the period which elapses before the cash is available in the United States. All these circumstances make the American exporter willing to sell a bill for £1000 for a less sum than par—less than $4866. The amount of reduction to which he will submit will be only such as offsets the total expense of sending to England for the specie. That expense is surprisingly small—between England and the United States, somewhere about one-half of one per cent. The bill of exchange for £1000 will not sell for less than $4845, or $4.845 to the pound. This is called the specie-importing point. When foreign exchange is at this point, specie begins to come in.

Under these conditions, *all* of the exporters' bills will be at a discount: all will sell for less than their par value. Competition being active between the exporters, no one of them will be able to sell his bill for a higher price than the others. The expense of shipping specie will have to be met by some one or other among them; to each one it is immaterial whether he will sell his bill at a discount or will send for specie. The market rate for all bills, when there is a continuing excess of exports, will be at the specie-importing point.

The reverse situation appears when the imports exceed the exports. The importers then need to buy more bills than the exporters can supply. Some of them will have to send out specie, and this involves the same sort of expense as bringing specie in. An importer who has to remit to London can afford to pay more than $4866 for a bill of £1000, rather than send specie. He will

pay as much, say, as $4885. If called on to pay more than $4885, he will refuse; for he can ship $4866 to England, and have this coin there converted into sovereigns.[1] Foreign exchange in New York will be at a premium, the extent of that premium being limited by all the expenses involved in transporting specie. The specie-exporting point, determined by these expenses, is about $4.885. When there is a clear excess of imports over exports, exchange will be at this premium; and, in like manner as in the other case, all the importers will have to pay this premium, even tho most of the transactions are liquidated thru bills.

§ 3. These are the simplest supposable conditions. They are rarely met in real life. Here, as in almost all the buying and selling of modern communities, a class of middlemen intervenes. The exporters and importers do not deal directly with each other; neither do they concern themselves with the possibilities of shipping specie in or out. They go to the dealers in foreign exchange, who are sometimes firms which make a specialty of this sort of business (the so-called foreign-exchange houses), sometimes banking firms which join this with other business. All the middlemen buy exchange constantly from the exporters and sell it constantly to the importers. They have their well-known correspondents in foreign countries, either branch houses of their own or other banking firms; they sell bills on these and meet bills drawn by them. When the exporters offer more bills than the importers will presumably take, the dealers none the less buy them; only, calculating that there will be no market for all the bills, and that some will have to be sent abroad and specie got with them, they will buy only at a discount. On the other hand, when the importers demand more bills than the exporters have to offer, the dealers sell to the importers, at a premium, whatever bills the latter want, and themselves send abroad the specie with which to meet these bills when presented. Being in the business and equipped for it,

[1] Foreign coin, or bullion, when it reached England, was in former days always taken to the Issue Department of the Bank of England, which was obliged by law (so long as specie payments were not suspended) to give notes for gold at a fixed rate, involving a very slight charge to the holder of the gold. The Bank of England thus acted as intermediary for the conversion of bullion and foreign coin into English money. It so acts still, after the abandonment of the gold standard in 1931; tho this part of its operations has ceased to be of importance.

they can ship specie more economically than the importers or exporters. Tho they make a profit, it is based on a very narrow margin.

With the presence of dealers comes that process of close bargaining, speculation, equalization, which naturally ensues with the specialization of trade. From the description just given of the simple case—that of exporters selling directly to importers—it might be inferred that if there was any discount at all or any premium at all, it would be up to the full limit set by the expense of shipping specie. But with the higgling and speculation among dealers, a discount of premium will appear which may be well within these limits. If, for example, more bills are offered by exporters at a given time than the importers are buying, the dealers may yet anticipate with confidence that before long a turn will come the other way and that at the later time the importers' demands will be in excess. They will buy the exporters' bills and wait for the turn. Possibly they will hold the bills in their own hands for a while; possibly they will send the bills to their foreign correspondents, tell these to collect the money from the foreign debtors, and hold the amounts until drawn against later. The current rates of interest on demand loans and short-time loans are important factors in these operations. If "money" is cheap (the rate of interest on bank loans is low) in the dealer's own country, he will more readily buy exporters' bills, and pay a better price for them. If money, again, is dear in the foreign country, he will also buy such bills more readily, since he can send them to the foreign country and there get a balance to his credit on which interest at a good rate is allowed. To figure out the price at which it is profitable to buy or sell exchange, calls for nice calculation of a number of items each involving a very small fraction—the direct expense of transportation, the mint charges and delays, the rates of interest in different countries, the probabilities of shifting currents of trade. Competition among the dealers leads to a market rate somewhere between the two specie points.

If, indeed, there is a continued balance of payments to be made one way or the other—if there is a steady and considerable excess of imports or of exports—then exchange will go to the shipping

point, and specie will flow in or out. The operations of dealers may enable the imports and exports to catch up with each other, and so may postpone the shipment of specie; but where there is continuing excess one way or the other, it moves in or out.

In the examples here chosen, we have spoken as if all the transactions in foreign exchange took place in New York—as if the London merchants were passive, and waited for those in New York to buy and sell exchange, and remit bills to London in settlement of the debts. In fact some of the transactions take place in each country. Which of the trading persons shall take the initiative in any particular case, depends on the bargain between them. It may be arranged that the New York exporter shall draw on his London customer, and so sell in New York exchange on London; or the London customer may assume an obligation to remit to this New York vendor, and so buy in London exchange on New York. Both sorts of transactions are going on all the time and in both centers exchange between London and New York is constantly being dealt in. When in New York English exchange is at a premium, then in London American exchange is at a discount. All the transactions are under the watchful eyes of the dealers; a remarkably close equalization of rates is brought about; while at the same time there is play for profit and speculation in terms of fractions of one per cent.

Bankers' bills, so-called—the bills drawn by dealers and bankers on their foreign correspondents—naturally sell for a somewhat higher price than most mercantile or trade bills. They contain the names of persons and firms well known in the business world. Again, sight bills naturally sell for a higher price than time bills. Foreign sales of merchandise, like domestic sales, are usually on time. The exporter who has sold goods is then entitled to receive his money at the end of thirty days, sixty days, or whatever the period for which credit is given. He draws his bill payable after thirty or sixty days, and discounts it at his bank. The bank, if it deals in foreign exchange itself, may hold the bill till maturity. Perhaps it sends the bill abroad at once to its foreign correspondent; perhaps sells it to a dealer in foreign exchange, at once or on maturity. The price at which it will sell depends on the length of

time it has to run, on the current rate of discount, on the calculations of the probable state of foreign exchange at its maturity. These minutiæ and others need not here be entered on. They do not affect the broad questions of principle regarding money, prices, and international trade with which we are chiefly concerned.

§ 4. The rates of foreign exchange are determined, not by the dealings between each separate pair of countries, but by those between a country and all the other countries with which it trades. The exports from the United States to England may much exceed the imports—in fact, they do greatly exceed them every year; but exchange none the less may be at par, if the United States imports heavily from other countries.

This situation is illustrated by the state of trade between the United States, England, and Brazil. The United States exports great quantities of cotton and foodstuffs to England; much greater in value than the manufactures which it imports from England. England exports manufactures to Brazil, greater in value than her imports from that country. Brazil, again, exports largely to the United States (chiefly coffee), but imports thence comparatively little. A merchant in New York who has bought coffee from one in Brazil could not easily find an American exporter who had bills of exchange on Rio de Janeiro or Bahia to sell. But he could find plenty of exporters who had sold grain and cotton in England, and had bills on London and Liverpool to sell. He buys English exchange and with this pays his debt in Brazil. Bills on London are welcome to the Brazilians, since in that country there are payments to be made for purchases of English goods. All these exchange transactions, of course, do not take place directly between exporters and importers, but thru the bankers, who buy and sell the bills and take keen advantage of every opportunity for equalizing payments without the shipment of specie. Thus by the mechanism of bills of exchange the exports of grain from the United States to England serve to pay for the imports of coffee from Brazil to the United States; and these same shipments of coffee, viewed as exports from Brazil, serve to pay for Brazil's imports of manufactures from England.

It does not much matter whether the bills which serve to settle

such cross payments are drawn on one country or another. Sterling exchange bills, drawn on London, were long the most widely used. Great Britain's enormous international trade ramified into all parts of the world. Many English banks and firms had well-established repute as dealers in foreign exchange. England had great industrial prestige; and the pound sterling was the best-known unit of value for the whole trading world. Hence foreign-exchange transactions were apt to be settled thru London and by bills drawn on London. During the war of 1914–18 this tradition was shaken, largely because gold could no longer move freely in and out of England, and a sterling bill in consequence no longer represented beyond question a fixed amount of gold. It is probable that in any event the growing international connections of other countries, and especially of the United States, would have brought about sooner or later some diversification of usage; the war hastened the tendency.

Much more attention is given to this detail in the mechanism of foreign trade than is warranted by its importance. Competition in the foreign exchange market is keen, the profit is kept within an extraordinarily narrow range, the business goes to those who do it most cheaply. But a certain patriotic nimbus attaches to these matters, as to most others connected with foreign trade. It disturbs the American or German or Frenchman that a sterling bill should have greater vogue than one in terms of his own currency. He is apt to think, too, that the use of sterling bills is part of cunning British schemes to divert trade from his own compatriots. There is no subject in economics to which more fallacies and prejudices attach than to foreign trade. Dollar exchange, franc exchange, mark exchange, are thought of not merely in terms of profit and loss but in terms of national glory. To the dispassionate observer it seems of small consequence which unit happens to be used or which set of banking houses do the business. True, the causes that lead to the choice—the character and extent of British trade or the growing importance of American—may mean much for a country's material welfare; but the designation of one or another place or country as that on which bills are customarily drawn is a matter of little moment.

Whatever the details in the mechanism of foreign exchange by which transactions are wound up, the total imports of a country are to be compared with its total exports. The state of the foreign exchanges—whether they shall be in general at a discount or at a premium—depends on the whole of its international trade. England, for example, exports manufactured goods to all parts of the world, and with these pays for her imports of foodstuffs and raw materials, of which she procures a large part from the United States. The United States, again, buys tea, coffee, cocoa, spices, sugar, jute, rubber, gums, and the like, from tropical and semi-tropical countries, and mainly pays for them, not by direct exports to these countries, but by exports of grain, meat products, and cotton, to England and other European countries. If there is a general excess of imports over exports, foreign exchange is at a premium, and specie tends to flow out; while a general excess of exports brings exchange to a discount, and causes an inflow of specie.

§ 5. Assume, now, that the total exports do not suffice to pay for the total imports (still assuming also that merchandise transactions are the only ones made and that the gold standard prevails). Then the imports must be paid for in specie. Will that specie flow out for an indefinite time? and what likelihood is there that a balance will permanently remain to be paid in this way?

The accepted answer to these questions, and in essentials the accurate one, was that the flow of specie sets in motion forces which sooner or later stop the flow. When specie leaves a country, prices tend to fall. Hence that country becomes one in which it is advantageous to buy; lower prices stimulate exports. Conversely, the country to which specie flows tends to have rising prices. It becomes one in which it is advantageous to sell; higher prices stimulate imports. But the flow of specie has an automatic limitation. The greater the flow is, the more certain is it likely to cease; the longer it has gone on, the sooner is it likely to cease. Merchandise exports and imports on the whole and in the long run balance in terms of money, because of the effect of the quantity of money on prices.

This is the answer in its simplest form; it is the statement of

the fundamental principle. But, like other economic principles, it holds true of the course of industry only in general. In details, it needs to be qualified and explained.

The modern mechanism of banking, currency, and international trade may be said to have had an innate repugnance to the flow of gold from country to country. All sorts of devices were resorted to in order to prevent or lessen such a flow.

Most familiar and effective among these devices is regulation thru the rate of discount. Gold, like any other form of money, is free capital, or command of capital goods; and it is, moreover, the kind of capital which is in every country equally available. It tends to go to the place where the return on loanable funds is largest. When specie first moves out of a country, it comes ordinarily from the bank reserves; and when it goes into a country, it goes ordinarily into the bank reserves. The rate of discount rises as bank holdings become less, and falls as they become greater. These changes of themselves tend to counteract the movement of specie in large quantities. The great central banks of England, Germany, France, Austria, and other countries, when they were operating under the gold standard, systematically raised or lowered their rates of discount in order to protect their specie holdings. The same thing happened, tho with less direct and conscious intent, under the banking practices which prevailed in New York before the establishment of the Federal Reserve system. Under this system, the method became the deliberately approved one in the United States also.

Such "protection" of a country's gold reserve was and is often carried beyond the point of reason. The advantage of a great hoard of gold and the harm of losing specie are commonly exaggerated, and this for various reasons. Something is due to the belief that a great stock of gold is a political or military asset. Something, no doubt, is due to the persistence of the older mercantilist notions, under which the specie supply was thought to be the most important constituent of a country's wealth. Not a little is ascribable to the repugnance of the business community to anything that tends to lower prices. Certain it is that an outflow of specie is usually regarded as damaging and is resisted by adjust-

ment of discount rates and like measures more strenuously than is warranted by its effect on a country's welfare. The flow of specie is an indicator of the currents of international trade, not in itself a matter of serious importance. And in the main it takes care of itself, ceasing by a quasi-automatic process when the needed readjustment of imports and exports has been accomplished.

To return to changes in the rate of discount. Often these affect not so much the volume or the flow of gold as the time when it takes place and its direction.

A rise in the rate brings an additional pressure to bear on those foreign-exchange dealers who may be preparing for a shipment of specie. Higher interest on money makes it more profitable to keep the money at home. It tempts bankers to wait; perhaps the currents of foreign trade will shift and enable the demand for exchange to be met without the shipment. Or it may lead such persons to arrange for shipment of specie from some other country. If reserves are low and discount rates high in England, and the contrary is the case in Germany, English bankers may buy exchange on Germany, and thereby secure the means of shipping specie from Germany to the United States. Very sharp calculations and very minute fractions in rates (both in rates of discount and rates of exchange) suffice to turn the currents one way or another.

Still another phase of international dealings is connected with changes in bank discount—movement of securities from one country to another. This is part of the general process of lending and borrowing between nations, of which more will be said in the next chapter. It suffices here to point out that the prices of securities in any one country are usually affected inversely to the rate of discount, rising as this falls, and falling as this rises. Hence securities which have an international market may be sent in place of specie in settling balances. There are brokerage firms which make it a business to watch the fluctuations of such securities in the different markets—London, Berlin, Paris, New York—and to buy in the one and sell in the other, on very slight margins of profit; and these dealings are closely dependent on the foreign exchange market and in turn are of prompt effect on that market.

None the less all devices, of whatever sort, do not prevent the

movement of gold or its ultimate effect on prices. They serve only to regulate and equalize it—to prevent it from taking place with a rush or from having sudden and rapidly disturbing effects. When there is a long-continued balance of payments in favor of a country, specie flows into it. Gold is constantly moving from country to country. Hardly a month passes without some shipment of it into or out of each of the important countries. When there is a balance of payments to be made to any one country because of a considerable and sustained excess of its exports, the current of gold not only moves in that direction, but continues to move until gradually equalization is brought about by changes in prices and by a restored balance between the countries.

Sometimes this result is reached without any movement of gold at all, or with a movement that seems not at all in proportion to the result. A country may be issuing paper money, or increasing its bank notes or deposits—operations which in themselves tend to expel gold. Then what happens is that the country retains both its paper and more or less of its gold, and gets rising prices without any influx of the metal. Again, the country may be one that mines gold. Ordinarily a mining country exports gold in the normal course of its international trade; but when its exports of other things are heavy, it may retain within its own borders the gold which would otherwise go out. The United States is an important gold-mining country, yet for several decades after the resumption of specie payment in 1879 kept within her borders the whole product of her mines; indeed, imported a substantial amount of gold in addition. The supply of specie which thus gradually accumulated was the result of a continuing excess of exports and was the basis of a tendency toward higher prices.

The consequence of all these modifying factors is that the flow of gold from country to country takes place, as a rule, not by large movements at any one time, but by driblets going sometimes one way, sometimes another; often by little-noticed diversions of the fresh supplies from mines. The comparative smallness of the ordinary flow is due mainly to the fact that international trade, long maintained, has already brought about such a distribution of the precious metals, and such a range of prices in the several coun-

tries, that their exchanges balance very closely. It is only when great economic changes occur that a large movement of specie takes place; and even then it is commonly distributed over a period of several years. The United States, exceptional in so many of its economic characteristics, long presented in this matter also the most marked exceptions to the usual situation. Not infrequently—as, for example, in 1879–80, immediately after the resumption of specie payments, and again in 1896–97, after the close of a severe period of depression—a great change in the relations between imports and exports caused sudden and heavy inflows of gold, giving the foundation for a rapid and sharp rise in prices.

So insignificant were the ordinary movements of gold from one country to another, so likely to be disguised by eddies and cross currents due to the complexity of international dealings, that some writers pooh-poohed the whole theory of the equalization of imports and exports by changes in international prices. Yet without this theory it was impossible to explain the facts, and especially the equalization of the money value of exports and imports. The influence of the quantity of gold on prices, slow in operation tho it was and subject to all sorts of disturbing elements, was the underlying persistent force which determined not only the international distribution of specie, but also, as will appear in the chapters that follow, the variations in the purchasing power of gold in different countries and the greater or less extent to which the several countries shared the gains from international trade.

CHAPTER 32

THE BALANCE OF INTERNATIONAL PAYMENTS

§ 1. Other items than merchandise exports and imports. Lending and borrowing and their effects on exports and imports. International dealings in securities.—§ 2. Expenses of travelers and non-residents. Remittances from the United States by immigrants. Freight charges.—§ 3. Position of a country that mines specie.—§ 4. Illustration from the international trade of the United States, 1790–1908.—§ 5. The notion of a favorable and unfavorable balance of trade. Usual attitude of the business community. In the main, an excess of imports or of exports is no indication of loss or gain; least of all, in the trade between one country and any other country.—§ 6. Unusual character of the international trade of the United States during the war of 1914–18, and again in the post-war period.

§ 1. In the preceding chapter, foreign trade has been treated as if imports and exports of merchandise were the only items in the balance of foreign payments. But there are other items, often of great importance. In analyzing them, I shall continue the procedure of the preceding chapter; that is, shall assume the trading countries to have one and the same monetary standard,—the gold standard.

The most considerable of the other items are loans—lending and borrowing between countries, usually thru the process of loans made by individuals or corporations in one country to like private entities in another. Loans contracted by governments are also common, the lenders being commonly individuals or corporations in foreign parts. Occasionally one government lends directly to another; huge transactions of this kind took place during the war of 1914–18. But like many of the economic events accompanying the great convulsion, these were of quite an unusual type.

There is a prevalent mode of speech concerning foreign borrowings, and indeed concerning foreign transactions generally, as if they always took place between nations as such—as if Germany as a body politic sold to France, or the assembled United

States borrowed from the English people at large. What happens in the usual and typical case is not that governments or peoples act, but that individuals in one country deal with individuals in another; in the case here under consideration, that individuals of one country lend to those of another, or possibly lend to the government of another. The usage of personifying countries, while often convenient for brevity, covers up and promotes a misunderstanding of the actual situation and sometimes foments unreasonable prejudice.

Suppose persons in the United States to borrow from persons in England. In the course of such transactions, entered on with a view to investment in the United States, the English lenders are usually bankers, upon whom American borrowers become entitled to draw for the sums lent. The Americans have bills on London to sell. If imports and exports have balanced before, there are now more bills on London offered in New York than the importers wish to buy. Foreign exchange falls in price and specie flows to the United States. If, indeed, these same borrowing Americans happen to make purchases in England—if, for example, they are railway projectors and buy rails at once in England (a common transaction in the second and third quarters of the nineteenth century)—then they may send their London bills directly to the rail makers in that city. In so far, the loan may be effected by the immediate import of commodities and without any flow of specie. But such a combination of borrowing and purchase is by no means universal. Ordinarily, the borrower wants money, or purchasing power; he may use his purchasing power at home, or in the lender's country, or in a third country. The loan is likely to bring in the first instance a fall in foreign exchange in the borrower's country, and a flow of specie into that country.[1]

[1] In the enormous loans (some ten billions of dollars) which the United States government made to its allies during the war years 1917–18, purchases in the lending country were linked with international advances in quite an exceptional way. It was stipulated, expressly or by implication, that the proceeds of the American government's loans should be spent in the United States itself. In fact, so far as the foreigners were concerned, no money passed at all. The Treasury opened "credits" on its books for the British, French, Italian, and other governments. The representatives of the several countries, as they bought goods in the United States, satisfied the vendors by giving them orders on the Treasury. No money was turned over to the foreigners at any stage. In effect, goods were turned over to them in

If this, however, goes on year after year, the same effect is produced on foreign trade as if there were an excess of imports into England. The flow of specie will not go on indefinitely. There will be changes in prices in England and the United States, such as will stimulate exports from England and imports into the United States. The imports into the United States will not necessarily be from England; there may be greater American purchases in third countries or greater English sales in third countries, or there may be both. The effect is likely to come gradually and almost insensibly, thru a little noticed diversion of the usual flow of specie, and thru changes in prices that are slight and seem on the surface due to other causes. But experience has abundantly proved that a continuing balance of this kind, like a continuing balance arising from merchandise transactions, is not liquidated in specie. It is settled by an increase of merchandise exports from the lending country. Such a country shows before long an excess of exports and this supplies the wherewithal for remittances to the borrowing country.

Transactions of this kind are not usually sporadic. They give rise to a steady flow of remittances, to which the merchandise exports and imports accommodate themselves. For a long time not only England but France and Germany also were lending countries. Such countries, in the earlier stages of lending, show an excess of merchandise exports over imports, and yet no steady discount on foreign exchange and no outflow of specie. The continuing loans are effected by the exportation of goods. The process is one of which neither the lending individuals nor the exporting merchants are conscious. The influence of specie flow and of changing prices is often gradual, silent, and little observed. Sometimes it is accompanied in the borrowers' country by rapidly rising prices, expanding credit, active business, speculation, general surface prosperity, and in the end a halt to the movement thru an industrial and financial crisis. In the United States crises in the nineteenth century were usually commonly accompanied during

exchange for their promises to pay, and the goods were at once exported. Such direct and immediate connection between loans in terms of money and exports of merchandise (both on a huge scale in this case) is rare; it belongs among the many anomalous phenomena of the war of 1914–18.

the period of incubation by heavy borrowing from abroad; at first an inflow of specie, then rising prices, and gradually an increase of imports.

Suppose now that the process of lending and borrowing has gone on for a long series of years. Then another factor enters, and in time the situation is reversed. The borrowers have to pay interest on their loans. As more and more loans are made, the annual interest charge swells. The principal of each loan is remitted once for all; but the interest charge to which it gives rise goes on year after year. In time the interest payments that must be made to the creditors' country will equal, and eventually they will exceed, the payments on account of any new loans which may continue to be made to the debtors' country. To this new situation the imports and exports will in time adapt themselves. The lending country which at the outset had an excess of exports will in the end have an excess of imports. From England loans were made thruout the nineteenth century to all parts of the world. The interest charges on the old loans gradually swelled until they came to exceed the amounts newly sent out on principal account. Hence the foreign trade of Great Britain showed a large excess of merchandise imports over exports—an excess, it is true, partly due to other causes, but mainly to this one.[1]

A converse transition takes place in the borrowers' country, from an excess of credit charges to one of debit charges—from an excess of merchandise imports to an excess of merchandise exports. The change may take place slowly and silently or may be accompanied by a financial crash. The turning point in the United States seems to have come with the crisis of 1873. As will be pointed out presently, the foreign trade of the United States changed in its character after that year; a previous excess of imports was replaced by an excess of exports. Tho the shift was not caused solely by the new relations of principal and interest in the international lending account, this seems to have been the dominant cause. It is not strange that the transition should be initiated

[1] The capital sums invested by the British people in other lands were put in 1909 at the enormous total of £2,700,000,000 ($13,500,000,000), and the amount annually payable to persons in Great Britain at £140,000,000 ($700,000,000). *Journal of the Royal Statistical Society*, September, 1909.

by a crisis and that the first phase of it should be a period of business depression.

The great lending operations of modern times take place chiefly thru the sale of securities. When governments borrow, they sell their evidences of debt. When loans are secured for private enterprises, corporate stocks and bonds are usually sold. The result of long-continued operations of this sort has been that certain securities have an international market, and pass freely from country to country. They are largely used in settlement of international balances and often obviate a flow of specie. Especially is this the case where a temporary balance of payments has to be met. Then the bankers thru whom bills of exchange are bought may send such securities instead of specie. On the other hand, these transactions sometimes cause independent disturbances of foreign trade, and so of banking and financial conditions. If the securities issued by a country's government or its corporations come to be distrusted, they are likely to be sent back to that country for sale, and then to cause a balance of specie to leave that country. Thus in the years from 1890 to 1896, when the contest over the gold and silver standards was going on in the United States, foreign holders of American securities became uneasy and sent the securities to the New York stock exchange for sale—a movement which contributed to the outflow of specie during those years, and added to the causes of public and private embarrassment.

§ 2. We turn now to transactions other than loans. One of the simplest is that of payments made to persons living or traveling in a foreign country. Americans who have permanently established themselves in Europe and American travelers spend large sums. What they spend is put at their command in Europe thru the mechanism of the foreign exchanges. In the first decade of the twentieth century these sums had mounted to hundreds of millions of dollars a year. Usually the Americans have letters of credit, which enable them to draw on bankers. Their drafts appear in the foreign exchange market precisely as do the drafts of foreigners who have exported goods to the United States. If the merchandise exports and imports of the United States just balanced, then these travelers' drafts would cause exchange on the United States to be

regularly at a discount in Europe, and specie would flow from the United States. But to this situation, as to that arising from loans and interest payments, the merchandise trade in time adjusts itself. The sums which Americans spend abroad are supplied by merchandise exports from the United States; for such transactions (supposing them to be the only ones of the non-merchandise kind) would bring about gradually and insensibly an excess of exports, in obedience to the same causes which would bring the exports and imports to a precise equality if these alone constituted international dealings. In the same way, British India has an excess of exports; partly because there are interest charges on loans of long standing made by Englishmen to the government of India and to private persons there, partly because there are in England many pensioners from the Indian service to whom the Indian government must make regular remittances.

A curious and important addition to the payments of this kind came in the United States, after about 1890, from the remittances of immigrants to their kinsfolk in the old countries. The immigrants newly arrived are frugal; it is the second generation that accepts the more liberal spending ways of the prosperous country. The newly arrived send a good part of their savings to relatives and friends at home, very largely for the purpose of enabling these also to emigrate to the land of plenty. Thereby again a debit item appears in the foreign exchange operations, which adds to the causes bringing an excess of merchandise exports. This item reached surprising size in the first decade of the present century; it was supposed to amount in each year to at least $200,000,000.

Freight charges on imports form another item of the same sort. If the merchandise between two countries is carried solely in the ships of one of them, this one will have in so far a balance due to it. Thus the foreign trade of the United States was long carried on chiefly in the vessels of other countries, England having the largest share. The citizens of the United States made remittances on freight account; and they would have had to make them by the shipment of specie if the exports just balanced the imports. England is in the contrary case. Her people are great owners of vessels, and are carriers the world over. By itself, this factor would

bring specie into England if her imports just balanced her exports. As a matter of fact, the remittances that must be made from other countries for freight take their place in the general balance of international payments. They also add to the causes which lead in England to an excess of merchandise imports.

§ 3. In modern times a country which produces gold is in a peculiar situation. If this be the only item (or the dominant item) over and above ordinary merchandise transactions, the country will have regularly an excess of merchandise imports, just as it would have if travelers' expenses or freight charges had to be remitted to it. But it will also have a regular outflow of specie; and therefore foreign exchange will be regularly at a premium. The specie is in this case an ordinary article of export, like wheat or cotton or any other commodity. But it goes out only when the state of the foreign exchanges is such as to warrant its shipment. In the usual case, where a country for one reason or another has an excess of merchandise imports, foreign exchange is not ordinarily either high or low; it reaches the shipping points only on the sporadic occasions of balances to be met. In a mining country, however, the state of the exchanges is normally such as to cause the exportation of specie. This was the case in the gold-mining colonies of Australia, especially Victoria, for many years after the gold discoveries of 1850; and it is still in the main the case in that region. It was the case in Mexico, long the greatest silver-mining country in the world, during the period when silver was specie on the same terms with gold. Since the universal adoption of the gold standard, and its introduction even into Mexico, silver has there become a commodity like any other, and the exchanges are reckoned on a gold basis. For the first decade after the Californian gold discoveries of 1850, the United States was in the same position as Australia.

Later experiences of the United States with regard to the domestic output of gold illustrate some of the irregularities of international trade, and show in what complex ways the underlying forces work out their results. During the Civil War the gold in circulation was driven out by the issue of paper money; and thereafter, until the resumption of specie payments in 1879, the

annual product of the mines was steadily exported, gold and silver being alike regular articles of export. After 1879 the United States accumulated an immense stock of gold. It did so (up to the time of the war of 1914–18) mainly by the simple retention within her own borders of the output of the domestic mines. In some years, not only this product was retained, but much gold flowed in from abroad in addition. Tho it happened in other years that not only the yield of the mines was exported but even more, yet on the whole the domestic product and some imported gold also accumulated in the country. The course of prices was affected by these movements, thru those processes whose slow, irregular, and half-concealed working has been explained in preceding chapters.

§ 4. A glance at the figures of the imports and exports of the United States during the hundred years from 1815 to 1915 shows that a striking change in the relation of the two items took place in 1873. Up to that date, imports had regularly exceeded the exports; after that date, exports regularly exceeded the imports. The excess of exports in the early years of the twentieth century was enormous; during the decade ending in 1908 the annual excess was $500,000,000. The reversal in 1873 is easily explicable from what has been stated in the preceding pages. During the first three quarters of the nineteenth century, the United States had been a borrowing country, and moreover had been in the early stages of borrowing. Steady recurrences of new loans more than balanced the accruing interest on old loans. Until 1860 the United States, in addition, had been a shipping and freight-carrying country, and its shipowners had been earning freights payable by persons in other countries. After 1873 the situation changed. Tho borrowing continued, sometimes on a great scale, the annual interest payable to foreigners on the whole offset the remittances into the country on capital account. Freight charges became payable no longer by foreigners to Americans but by Americans to foreigners, for the reason mainly that iron steamers displaced wooden sailing vessels and could be built and operated more cheaply by the British and by others in Europe. Moreover, other debit items appeared newly, while others of the old kind rose to dimensions so much greater as to make them substantially new. The traveling expenses of Ameri-

cans became vastly larger; so did the remittances of immigrants. In some years, repayments of old loans were made, in the form of purchases by Americans of securities which in previous times had been sold abroad. Hence the preponderance of exports after 1873. At first comparatively slight, it reached eventually the remarkable volume just mentioned,—half a billion of dollars a year. Irregular tho the merchandise balances were, influenced as they necessarily were by the accidents of the season and the crops, by monetary legislation, by crises and depressions and "booms," the general trend was unmistakable; the exports advanced more rapidly than the imports, exceeded them in almost every single year, and in most years exceeded them immensely. The flow of specie meanwhile was at some times into the United States, in other times out of the United States. The upshot was an international redistribution of gold different from that usually occurring when a country mines the metal. The United States, so far from exporting its product, retained virtually the whole within its own bounds.[1]

§ 5. When the merchandise exports of a country exceed its imports, the country is said to have a "favorable balance" of trade. When its imports exceed its exports, the balance is said to be "unfavorable." The same terms are used when the state of international trade is such as to cause an inflow or outflow of specie; altho, as we have seen, such inflow or outflow is by no means a necessary or even a usual consequence of an excess or deficiency of exports. The general notion underlying these terms is that a country gains by having dealings with other countries which are expected to bring specie in, and loses by those which are expected to take specie out.

This notion goes back to the Mercantilist writers of the seventeenth and eighteenth centuries, who believed that specie was a peculiarly important part of a country's wealth and that legislation on international trade should be directed to its accumulation. Any one who has grasped the elementary truths about wealth, exchange, money, will see the absurdity of supposing that the prosperity of a country is bound up with the inflow or outflow of

[1] A history of the relation between imports and exports in the United States during the pre-war period is given in the Harvard *Review of Economic Statistics*, July, 1919.

specie. The astonishing thing is that, notwithstanding the simplicity of those truths and their repeated and widespread exposition, ignorance regarding them should be so common. Many people who think themselves entitled to attention still speak as if an excess of exports promised a profit to a country, and an influx of specie were a realization of that profit.

In part this way of looking at international trade comes from the habitual attitude of business men and bankers. Plentiful bank reserves, low rates of discount, easy accommodation of borrowers —these are always welcome to the business community. Conversely, diminishing bank reserves and higher discount are unwelcome. Hence the inflow of specie, which proximately affects bank holdings and short-time interest, is spoken of as a good thing and the outflow of specie as a bad thing. The outflow, with its consequent pressure on loans, interest rates, and eventually on prices, is often a salutary check on inflation and speculation. But few business men feel it to be salutary. Nearly all would like to see an unending round of rising prices.

There are times, of course, when the balance of international payments—usually resting on the relation between exports and imports—is of real consequence. This is notably the case when a country is trying to extricate itself from a depreciated paper currency. The return to specie payments at the start of the operation is possible for such a country only if its foreign trade is in a state which will cause specie to flow in, or will prevent it from flowing out when a fund for resumption purposes has been collected by the government. Tho in the long run this inflow or outflow will depend on the state of prices, in any one year or even period of years the balance of international payments is affected by all the happenings of the time. If at the time when a country is preparing to return to a specie basis, financial disturbances or poor crops lead to an "unfavorable" balance, the operation of resumption will be difficult and perhaps unsuccessful. Conversely, a favorable balance may make success easy. It was an immense aid to the resumption of specie payments by the United States that in the year fixed for it (1879), and in the years immediately following, there were unusually heavy exports, because of good crops within the country

and poor crops elsewhere; while at the same period improvements in railway transportation greatly facilitated the outflow of exports. The consequent inflow of gold, coupled with the retention of the domestic output of the metal, gave an unexpectedly solid basis to the reëstablished specie régime.

In the main, however, the persistence of the Mercantilist attitude is not due to any such considerations but to simple ignorance. Since people measure their individual incomes in terms of money and their profit by an excess of money receipts over money expenses, they fall easily into the way of regarding money as the important form of wealth. This was indeed the earliest and crudest form of the Mercantilist notion. It is the same ignorance and fallacy that underlie the advocacy of paper money inflation and the various schemes for making nations prosperous by adding to their stocks of cash.

One curious form of the Mercantilist view appears in the interpretation often given to the state of trade between a country and a single one of its neighbors. Thus the exports from Canada to the United States may be greater than the exports from the United States to Canada; and it is often inferred (for example, in discussion of reciprocity treaties between the two countries) that the trade is one unfavorable or damaging to the United States. Such comparison is meaningless. So far as the relation between imports and exports is a matter of moment at all, this is to be judged by the balance of transactions not between any one country and a single other but by its balance with all. That our exports to England exceed our imports thence, and our imports from Brazil exceed our exports thither—all this signifies nothing. It must be confessed that public men in high station, as well as newspaper scribblers and rabid partisans, fall into loose talk on this subject, and compare the sales and purchases of one pair of countries as if these really gave an indication of their relative gains from trade with each other.

§ 6. During the decades immediately after the end of the war of 1914–18, the course of international trade and international payments illustrated the wide variety of factors that enter. It cannot be said that anything appeared which serves to impugn the analysis

and reasoning which have been set forth in the preceding pages. But the complexities were extraordinary, and problems of a new kind emerged,—new not as regards economic principle, but in the field of economic policy.

First of all, an economic factor: The United States became what is called a "creditor" nation. By this term, often vaguely used, I mean what in general it implies, namely that the people of the United States not only had lent to people in other countries but that the stage had been reached where the sums due to them on interest account overbalanced what would run the other way if there were further new loans. In Great Britain a marked creditor position had come long before, thru a gradual process extending over a couple of generations; in the United States the war of 1914–18 brought it about with dramatic suddenness.

During the war itself, the overturn was forced by the British themselves thru a process till then unheard of. Their government virtually laid hands on the American securities held by individuals in their country (paying the holders a fair price) and sold the securities in New York. The proceeds were then used in New York for buying American goods wanted for the conduct of the war. American securities having a value something like a billion dollars were thus taken over and sold to Americans, while goods to the same amount were bought by the British government. It was a case unusual in two ways: the British government deliberately redirected private investment; and also (what was less unusual) great changes in merchandise movements took place without any transfer of money at all.

A second series of transactions, in some ways similar but in others quite different, was on an even greater scale, indeed on a scale quite beyond precedent. After the United States entered the war in 1917, the American government lent to its allies enormous sums, some ten billions of dollars. The loans were effected by putting at the disposal of the agents of the allied governments in the United States (quite the most important agent being the great Morgan banking firm) billions which were then used by these agents within the country for purchases of supplies wanted by the allies. The goods thus bought moved out of the country without

any corresponding inflow either of goods or specie, or even of securities. By the first of the processes (that described in the preceding paragraph) the British private creditors of American debtors were replaced by American creditors, and so far (tho by no means as regards the whole world) Great Britain ceased to be a creditor country. By this second, the British government became directly a debtor to the American government, and Great Britain became formally a debtor country.

Some years after the close of the war, international lending and the change to a creditor position took place in the more usual way. Large loans were made by American citizens to foreigners, partly to governments, partly to private concerns. How the American bankers who managed these operations got command of the very large funds thus loaned by them, or thru them, is not easy to state summarily. In part, no doubt, they represented true savings. In part, they rested on "created" money; bank deposits which were created in this case the more freely because the great development of the Federal Reserve system made the growth of loans and of deposit money more easy than ever before. At the stage of transition in which the whole banking system then was, the available banking funds seemed almost limitless. Millions were lent to borrowers, not only in European countries but in South America, Australia, Canada, all over the world. At a still later stage American banks lent on a great scale for short periods—short-time loans, made not with expectation of any sales of securities to foreign investors but merely for the purpose of finding something to do with the sums which accumulated in the banks —still in the main thru the creation of deposits fostered by the Reserve System. Partly as cause and partly as effect must be reckoned the speculative mania of the period, the like of which in quantitative terms had never before been seen.

When the whole top-heavy structure came to the ground with a crash in 1929, not only did out-going loans come to an end, but interest on many of those already made ceased to be paid. Many debtors became bankrupt. International payments, and so international trade, entered on a period of uncertainty, haggling, shattered mechanism, distress on all sides.

To these complications in the relations between private leaders and borrowers was added another, arising from the interallied debts—the enormous loans made by the United States government directly to the governments not only of Great Britain but also to others of its former allies. The chapter is a dreary one, beclouded by bitter controversies. As regards the consequences for international trade, it is enough to note that, while payments of interest and repayments on capital account were made for a while, they ceased soon after the collapse of 1929. In these relations, as in many of the private ones of the time, the new "creditor position" of the United States came to be no more than nominal. The loans from government to government, which led to such enormous exports when made, never had any corresponding reverse effects on imports, as they would doubtless have had if repayments had been made.

Concurrently with all the old and new interest and capital payments, the other "invisible" transactions went on. Imports and exports, the visible items, went their course, less regularly than in ordinary times, yet in far more familiar fashion than the nonmerchandise operations. Somehow, as the unexpected resultant of the whole complex, there ensued a great flow of gold into the United States. It is easy to understand that such a flow should have taken place during the war itself, when there was the sudden insistent demand for American goods. But its continuance (even tho occasionally offset by a return movement) on so great a scale for so long a time is one of the most striking among the anomalies which appeared during the decades following the war.

One further aspect of the anomalies of the post-war period is to be noted: the increase in the movement of securities from country to country and the new and growing pressures from the security flows. That these transactions were greater in volume than ever before was not mainly the result of more lending and borrowing by individuals for investment or by governments for immediate expenditure; it arose mainly from irregular movements to and fro, especially from "flights of capital." The disturbances and convulsions, both economic and political, in European countries led large and small investors to shift their holdings first one

way then another; and there were times when American investors did the same, taking refuge in European (especially British) securities. An ominous prospect would cause a rush of sales at home and purchases abroad, and movements of gold between countries in the course of these operations; then return movements as the skies cleared. The transactions led to gold flows which impinged on the banks, and thereby with more or less rapidity and more or less real effect on the business situation. It was difficult to perceive what would be their consequences, quite impossible to predict their occurrence.

This connection between gold movements and international security movements is of a different kind from that noted earlier in this chapter. It meant not that security sales were a substitute for gold movements, serving to make the gold flow less; it meant that the security movements, induced by extraneous factors, operated to make the gold movements greater and more irregular. Under such conditions, the banks of a country, and especially central banks, are suddenly and unexpectedly called on to provide large sums of gold for remittances to foreign countries. If their holdings are slender, they may be led to shut down on specie payments once for all; or if not to act so drastically, at least raise their rates of discount suddenly and sharply. Even if their specie holdings are adequate, they may tighten the reins and bring a halt in the world of banking and business. In this sort of situation, as in so many others, the psychological element plays a large part. The supposed disadvantages from an unfavorable balance of trade and an outflow of specie and the supposed advantages of the contrary are so constantly harped on in the financial press and in the everyday talk of business men and bankers that there is always an uneasy feeling when gold flows out; and this feeling easily becomes panicky when the outflow is large. Certain it is that the remittances occasioned by security transactions became both greater and more irregular in the post-war period, and affected more and more international payments and the play of the foreign exchanges. Not least, they affected the smooth working of the gold standard, and were no small factor in leading to the doubts about its retention.

On the matters taken up in this chapter, and more particularly on the international trade of the United States as it had been affected by the war conditions, I wrote in 1927: [1]

> The export of capital may wax and may wane. It will hardly cease; but it may prove of less commanding importance than commonly expected. The great excess of merchandise exports over imports may continue, or may be succeeded by an even balance, or may be replaced by an opposite relation—an excess of imports. The whole balance of international payments, with all its interwoven and conflicting elements, subject as it is to the possible decline or growth of almost every item, may take a shape quite beyond any foreshadowing. Most interesting of all, and perhaps most unpredictable of all, is the way in which the monetary mechanism will operate. Will the currency system of the country remain as insensitive to gold movements as it came to be during the war and post-war periods? Will the whole enormous gold stock be retained within the country? And if not, what will be the method and what the results of its redistribution? Will deliberate policy and plans for a managed currency determine its movements, or will its international distribution proceed after all in that undesigned fashion which has so long been regarded as part of the normal economic order? . . . Just how the prices of American goods, and the country's general level of domestic prices and money incomes, will shape themselves or will be shaped under the newly developed banking and currency situation—he would be a rash man who ventured to predict. And hence he would also be rash who ventured on a prediction as to the ultimate structure of the country's international trade, of its merchandise exports and imports, of its balance of payments with other countries, of the greater or less extent of its real gain from the total of all the transactions taken together.

Twelve years later, at the time (1939) the present pages are written, the same questions are before us, and the same uncertainties remain about the answers. And it also remains true that the dispassionate student will be less interested in the shifts and changes that take place from year to year, even those from decade to decade, than in the play of permanent forces and the way in which these act on the real gain of the world and of the several countries from international trade. The world would be much better if trade between nations was carried on as it was in the two

[1] See pp. 333–334 of my *International Trade*, published in 1927.

decades that preceded the war of 1914–18 rather than in the way of the two succeeding decades. The gold standard then smoothed the international exchanges, and the economic structures of the several countries were adapted to monetary and industrial relations that worked probably as well as human institutions are likely to. The time may come when nations will again look at each other as friendly competitors, not as hostile combatants. Then more attention will again be given to the abiding factors that act on the trade between them. Unless the fundamentals are understood, it is impossible to render judgment intelligently about the past or discern what is good in the future. Let us hope that the fundamental matters that are forgotten in the welter of fencing on the immediate questions of money, banking, balances of trade, dislocated exchanges, will again come to the fore. In the next chapter one more of the short-time or "abnormal" factors will be taken up; in the chapters coming thereafter, we turn to the more essential things.

CHAPTER 33

DISLOCATED EXCHANGES

§ 1. Exchanges become dislocated when the currencies of the trading countries are not on the same basis, e.g. not on the gold basis.—§ 2. Effects of dislocation in the simplest case: when merchandise transactions are the only ones between countries, and prices in each country are stable in terms of its own currency.—§ 3. The case usually is complicated. Illustration: changes in tourist expenditures; effects on imports and exports of goods; "bounties" on imports or exports.—§ 4. Prices of domestic and of international goods do not move together.—§ 5. Dislocations arising from changes in the monetary standard itself. The United States in 1933–36.—§ 6. Effects of complete abandonment of the gold standard. "Exchange dumping."—§ 7. The effects of dislocated exchanges are of the short-period kind. In the long run the character of a country's international trade is not altered by them.

§ 1. In the preceding chapter international trade has been considered mainly on the assumption that one and the same metal —gold—is the basis of the circulating medium of the trading countries, and moves freely between them. It is obvious, however, that the machinery of the flow of specie, of prices, of money incomes rising or falling because of the specie movement, of readjustment thru a new level of prices and money incomes in each of the countries—all this cannot operate, certainly not all of it can operate, where the monetary systems rest on different bases. Where there is paper in one country and gold in another, or gold in one and silver in another, or paper in both, the mechanism is different. The compensating, correcting, stabilizing influence of gold movements ceases. The present chapter takes up the more complicated case where there is trade between countries having different monetary standards and the correcting influence of the gold standard is lacking.

The most obvious consequence of the absence of the stabilizing effect of a common standard—which in modern times meant the gold standard—is that foreign exchange becomes dislocated (to use the term I have chosen). While the range within which the

price of foreign bills of exchange can fluctuate is very narrow when the trading countries are all on the gold standard, it is wide when they are not. Exchange is out of gear; the price which is yielded by the command or possession of money in one country over the money of other countries may fluctuate with no limits accurately assignable. Since almost all transactions between foreign countries are carried on thru bills of exchange, fluctuations in the prices of bills affect the substantive course of trade—the imports and exports—in ways which do not appear where the same standard prevails in both. And herein, it may be noted, lies the most striking difference between domestic trade and international trade. Within a country there is no problem of dislocated exchanges; all transactions are on the same monetary basis. Not so between countries each of which has its own quite unique monetary basis.

§ 2. The range of fluctuations in dislocated foreign exchanges would be less and the consequent uncertainties about the outcome of individual transactions diminished if all transactions were carried on steadily in the same volume and at the same speed. The situation would be simplest of all if the only items in international payments were the imports and exports of merchandise, and if these went to and fro in a settled and unchanging routine. Given this situation, the rates of foreign exchange would be adjusted, nay could be said to adjust themselves, to the relative purchasing power of the several currencies in their own borders. Suppose, for example, there were two trading countries, say the United States and Germany, the United States having paper dollars and Germany paper marks (dismiss all thought of the gold dollars and gold marks with which the names of these monetary units are historically associated). Then if prices of goods in terms of marks in Germany were ten times as high as the prices of goods in the United States in terms of dollars, foreign exchange would be at the rate of 10 to 1. A bill of exchange on Germany for 10 marks would sell in the United States for 1 dollar, and a bill for $10 on the United States would sell in Germany for 100 marks. Within each country the price level would be determined by the factors there at work,—the volume of the paper marks and the

bank deposits and subsidiary coin, the volume of transactions, the rapidity of circulation of money and goods, and so on. Given that price level steadily maintained, and given the simple case of unchanging international transactions restricted to merchandise, the rates of foreign exchange would follow, and would be stable so long as these conditions remained unchanged.

No such simple situation has ever existed or, it may be safely predicted, ever will exist. But like other simple situations imagined by the economists, it serves well as a first step in analysis, as a clue to be kept in mind when watching the deviations from simple cases which appear in the world of affairs. And not only this: the facts show in the long run a substantial conformity to the theorem. Conformity of this kind is to be expected when we recall that merchandise operations remain far and away the largest constituent in the total range of international transactions, and that it is these which are most likely to go on in a steady stream. Further is it to be remarked that paper money systems, however easy it is to manipulate them, do attain not infrequently a certain stability. Whether the paper is put out from the start as "fiat money" or is a degeneration from a specie standard into that form, the volume is often maintained for considerable periods at roughly the same level. When once the initial stage of sudden great resort to inconvertible paper money has been passed,—when the first throes of financial stress are overcome,—a stage of comparative quiescence may ensue. Such was the case in the United States after the Civil War of 1861–65, in Austria after 1866, in France after the war of 1914–18. Given conditions of this kind, foreign exchanges show a similar sort of approximate stability. The rates fluctuate, but not very widely from a mean that corresponds to the mean of the price levels. So far as there is this correspondence, international trade goes on very much as if both countries had the same monetary standard.

§ 3. The situation is far from being as simple as this when there are changes in one or another of the factors, especially if the changes are abrupt. The process of accommodation to new price levels and to new rates of foreign exchange then takes place slowly and is trying. And even tho the monetary factors may remain

undisturbed—gold or silver used or not used in the same way as before, paper money not managed or manipulated on a new scale, the resort to credit instruments not modified,—the various non-merchandise items entering into the international accounts, such as capital investments or tourist expenditures, may fluctuate with unexpected and disturbing results. All such operations impinge in the first instance on the rates of exchange, then on the prices of goods, then by slow steps on the long-period or normal adjustment of the exchange of goods and services between nations.

It will serve to make more clear the nature of the problems that arise, the way in which the economic forces operate, if some illustrative cases are taken up.

First, suppose that there is a change in one of the non-merchandise items. The volume of tourist expenditures changes, and additional payments have to be made on that account. Suppose the change to be abrupt and considerable. Travellers from the United States, say, go in large numbers to Germany, spend freely there and need German money.[1] Their bankers supply them with exchange on Germany in terms of marks, or (what comes to the same thing) authorize them to sell in Germany, for marks, bills drawn on the American bankers, i.e. exchange on the United States. There is an increased demand in the United States for mark exchange, an increased supply in Germany of dollar exchange. American exchange would sell at a lower price in marks than before. To revert to the suppositious figures just used, a bill on Germany would sell in the United States not at the rate of $10 for 100 marks but for more—say $12. A bill on the United States in Germany would sell not at the rate of 100 marks for $10, but for $8⅓. This is the first impact: exchange would be at a different figure, and one not in accord with purchasing parity.

Consider next how this would bear on *other* transactions, and especially on merchandise imports and exports. The merchant in New York who has goods sold or salable in Germany, and has exchange on Germany to offer in the market, now finds that a

[1] I assume here what may be called the natural or ordinary course of events; not that complicated system of controls and regulations which was set up by the government after 1935 and which is as difficult to understand as it is (1939) difficult to enforce.

bill for 10,000 marks sells for $1200 instead of $1000. It pays to export; those who have goods on hand, salable at once in Germany, get a "windfall." Something in the nature of a bounty on exports results. It must be said at once that this effect on the prices of exported goods is not so certain and predictable as is often stated or assumed. It may be considerable, may be moderate, may be nothing or next to nothing. Take as example cotton. The immediate effect—still assuming the new demand for German exchange to be abrupt—would be to cause American cotton to go up in price. But soon it would become a question whether the added American cotton offered in Germany would sell there at the same price in marks as before. Unless indeed the quantity were negligible, the price would be lowered, the elasticity of demand in Germany determining the outcome. It would become a question too how far the Americans at home would lessen their own purchases of cotton. To the degree that *their* elasticity of demand led to smaller purchases, the amount of goods exported and offered for sale in Germany would increase and the price in Germany would be more considerably depressed. Elasticity of demand in Germany might be slight, and the additional American cotton salable there only with a marked fall in price; while the elasticity of demand in the United States might be great, and the decline in American consumption rapid. Then the effect would be not that the American holders of cotton got a bounty but the German purchasers a windfall. The "bounty" on exports thus is by no means certain to ensue. One would expect on the whole neither the full bounty nor the great windfall; a somewhat higher price in the American market, a somewhat lower price in the German. The actual course of events has usually been of this intermediate kind.

This sort of case can be generalized. We may suppose the tourist travel from the United States to *all* countries to swell, and all American exported goods to be affected; evidently with complications as regards different goods but a general outcome such as has been indicated.

It is worth while to remind the reader that the case can be generalized in still another way. The same *sort* of consequence ensues even if both countries (or all countries) have the same monetary

standard,—gold in the modern world. Exchange is affected in the same way, exports stimulated in the same way. The difference, a most important one, is that under a common gold standard the effect is slight. Exchange can fluctuate but little, and the consequences of fluctuations, tho the same in kind, are as a rule negligible in amount. It is the circumstance that dislocated foreign exchanges may fluctuate quickly and widely that makes the consequences of moment and leads to acute economic and political disturbances.

§ 4. Consequences of this kind, in all their various forms, rest on the same underlying cause: the fact that the prices of imported and exported goods—international goods—are specially interlaced with the rates of exchange. This is but one form of that dispersion of price movements to which attention has already been called.[1] Prices of international goods are subject to influences of their own. Their prices, while not unrelated to those of domestic goods, show movements of their own. The deviations in direction or degree may be small or large, may last for months or for years. They are affected by all the varying influences to which prices of particular goods are subject, such as changes in the arts of production, changes in demand, seasonal changes in crops, *and* to the conditions of international payments and international trade. In the case here supposed the American prices of export goods would go up, while there would be nothing to affect the domestic price level. The prices of goods made and sold at home, and the current rates of money wages, would remain as before.

In course of time, this discrepancy between the two price levels —of domestic and international goods—would presumably be evened out. The higher prices of the exported goods would tend to draw people to produce more of them, and they would come to sell in the same proportion to their costs as domestic goods, i.e. presumably at the same price as before. But the process of equalization is slow, especially in the case of agricultural commodities, such as cotton in our example. And during the period of transition, other factors are likely to enter, either working in the same direction or in opposite directions. The eventual situation then

[1] Chapter 20, on Changes in Prices.

is the combined effect of several causes acting at once, and the effect of no single cause stands out as separate.

The same sort of analysis can be applied to other cases. Anything which disturbs dislocated foreign exchanges has its repercussions on the prices of international goods and on their movement from country to country. The story is essentially the same whether there be investment transactions, interest payments, freight accounts or political remittances like those which were expected (and for a while were made) after the war of 1914–18. All these have their immediate effects and their ultimate effects. The ultimate effects are not essentially different from those which would come about if there were stable foreign exchanges,—the gold standard all around. But the immediate effects under the sensitive and quick-moving dislocated exchanges are not the same, and the processes of transition to the ultimate effects are more devious and take time,—often very much time.

§ 5. There is another class of cases, analogous to those just considered. It is that of a change in the monetary standard itself and a consequent dislocation of foreign exchange. There may be "devaluation," a modern term for a new form of an ancient practice. By this the amount of specie (typically gold, in the standard coin or unit of value) is lessened, the gold standard being retained at what may be described as a lower level. Or the specie standard may be given up, for a time or for good, resort being made to inconvertible paper. The effects of these two variants on the foreign exchanges and on international trade may be taken up in turn.

The remarkable experiences of the United States in 1933–35 illustrate the case of devaluation. In 1933, as the first step in a series of drastic changes, the United States suddenly broke loose from the gold standard. Gold was no longer received for free coinage at the mint, and paper money was no longer exchangeable for gold; nay, private persons or corporations were no longer allowed to hold gold, still less to use it as money. The standard money became inconvertible paper. But soon a hybrid developed, and the case became one of devaluation. First the Treasury announced that it would buy and sell gold at a price (in terms of "dollars") announced for the time being but subject to change from time to

time. Later, early in 1935 (February) the Treasury settled on a price of $35.00 per ounce of gold; whereas before 1933 the corresponding price—in effect a "mint price" tho not so stated in legal terms—had been $20.67. Under the law the President had been empowered to make changes from this long-standing figure at his discretion, and now put it, until further action, at the new figure. Devaluation, tho not settled definitely, was put into effect with no indication of further change to come. The period during which the process went on was nearly two years,—from April, 1933 to February, 1935; waveringly in the first stages but with a succession of steps from October 1934 until the final (?) "mint price" of February, 1935 was reached.

The object of this series of operations was to bring about a rise in prices all around. The mistaken view on the relation between gold supply and the general price level on which the operations rested, and their failure to achieve the expected rise in prices of commodities at large, have already been pointed out.[1] But there was a quick effect on the foreign exchanges and on the prices of imported and exported goods, these moving in the direction desired, even if not to the full extent desired. The government dollar prices for gold determined the price in dollars of bills of exchange on countries which still held to the gold standard, such as France, Switzerland, Holland. An American merchant who sold goods in such a country, or any one having a bank account to his credit there, could draw a bill in terms of gold and sell it in New York at a price corresponding to the government price. Exchange on countries which were not on the gold standard (Great Britain of course the most important) tended to move in the same direction. Necessarily the changes in these other rates were affected not only by the conditions generally prevailing but also by conditions peculiar to each one. Sterling exchange notably had a course of its own, being much regulated by the British Exchange Equalization Account. The intricacies of the happenings were not easy to follow, but the general trend was clear. Here too there developed something in the nature of a bounty on exports and a tax on imports. Prices of international goods, tied as they are to the rates of

[1] See Chapter 30, especially §§4–5, pp. 428–430.

foreign exchange, tended to rise in the United States. Domestic prices went their own way, affected little if at all by the devaluating operations. Those international commodities for which the export market was important as compared with the domestic market, such as cotton, were influenced most markedly, while those for which the domestic market bulked larger felt it less. As has been already remarked, the range of possibilities is wide. In this particular case most of all unexpected was a great seasonal disturbance,—a serious drouth in the agricultural West, a great crop shortage there, not counterbalanced by large crops in the rest of the world, and a sharp rise in the price of wheat and of meat products the world over. This coincided with the monetary operations and made them seem of much greater consequence than in fact they were. Hardly ever is it possible to make out with accuracy what quantitative effect is to be ascribed to any one of several economic forces which act together; in this case the crop vicissitudes were clearly quite the most important.

§ 6. The same tendencies appear when the specie (gold) standard is given up entirely. But the complications and uncertainties are greater, and they vary according to the way in which inconvertible paper is put out. Sometimes the issue takes place quietly and inconspicuously, as in the United States during the early stages of the Civil War (1861–62) or in France during the early stages of the Franco-Prussian war of 1870–71; sometimes at a rapid pace and on a huge scale, as in the disastrous German inflation of 1923 (this, to be sure, an aggravation of a comparatively moderate fiat money situation already existing). When the process begins slowly, it may be soon checked, as in the French case, and leave little trace behind; or it may quicken and enlarge, as in the more typical experience of the United States, and so may have consequences of great moment and long duration. What happens to the prices of goods and to money incomes has been already discussed; here we have to consider the effects on the foreign exchanges and thereby on international trade.

The mere fact of "going off gold" does not in itself necessarily affect the rates of foreign exchange. But bankers, speculators, middlemen of all kinds are at once affected by new expectations and

calculations; and it is their doings which have immediate and direct effect on the foreign exchanges. The first step in the suspension of gold payments is likely to be regarded as the prelude of more to come. The bankers and brokers who deal in bills and the exporters and importers and other customers of the bankers become uneasy. A premium on gold appears at once and a higher price for bills on a gold country. If issues of paper increase, and are expected to increase more, the tendency to rise becomes accentuated. While the ultimate and important effect is on the general range of domestic prices and incomes, the more immediate effect is on the prices of the imported and exported goods,—the international commodities. If, as so often happens, the process of currency inflation and of depreciation is prolonged over a period of time, there is during that period a rise in the price of exported goods; or, if the domestic price happens to be held constant, a fall in the foreign price of those goods; more commonly, the sort of intermediate effect already described. The tendency, and sometimes the marked and unmistakable tendency, is for the same kind of bounty on exports as has been described in the preceding pages.

That bounty arises, as in the other cases, from the *process* of depreciation. It ceases if and when the process comes to an end and stability is restored. Complete stability of prices of course is never achieved, any more than it is under the gold standard; but some approach to stability is sometimes reached. Foreign exchange none the less remains dislocated as long as the paper money is inconvertible, and international transactions of all kinds continue to have an unstable and speculative element. There is always the possibility that the process of depreciation will be resumed; and during the chaotic period of the third decade of the twentieth century it was sometimes resumed for the very purpose of restoring or continuing the bounty effect. The foreign exchanges became highly sensitive both to the expected and the unexpected fluctuations in the several items that enter into international payments; they were easily influenced by purely speculative operations; and not least were influenced by deliberate government action designed to affect the rate of exchange,—sometimes to prop it or to bring it up, sometimes to push it down. Many who write on

this subject have a notion that the mere existence of dislocated exchanges, of whatever sort and however prolonged, leads in itself to a bounty on exports, and that a country resorting to inconvertible paper thereby secures *ipso facto* a "more favorable" balance of trade,—a lessening of imports, an increase of exports. And this belief, coupled with the persistence of the mercantilist notion that a rise in exports is the one thing to be sought in international trade, led to sad results during the great depression after 1929. Nations were suspected of resorting to "exchange dumping" deliberately for the purpose of getting the better of other countries. National prejudices and hatreds were fomented. "Counter-devaluation" was demanded, or "defense" by protectionist legislation of the extremest kind. The monetary confusion led to evils reaching far beyond what had ever before been experienced from that source: a great obstruction of substantial international trade and a great decline in the physical volume of imports and exports.

To summarize. A long-standing dislocation of the international exchanges does not of itself have disturbing effects on international trade. If monetary conditions settle down; if prices as a whole accommodate themselves to a given standard,—then the rates of foreign exchange come into accord with the price levels of the trading countries and remain at figures which represent those price levels.

The substantive trade comes in the end to be carried on in the same way as if the trading countries had the gold standard (or any common standard). The general principles applicable to international trade remain applicable. What these principles are will be taken up shortly; it is enough here to indicate their tenor. Differences of comparative advantage operate in the same way: a country tends to export the things which are comparatively low in price within its borders; that is, those produced under advantageous conditions, in general at low real cost. Changes in the play of international demand, in capital investment transactions, in payments for services, have effects which are in substance the same as when foreign exchange is a stable and simple matter. The different working of the exchange machinery does have effects of its own, but these do not alter the essentials.

§ 7. The effects of dislocated exchanges on the substance of international trade and on domestic industry, then, are of the short-period kind. But in this case as in so many others the short period is not very short,—indeed, is short only in relation to long periods which are very long. The effects persist for greater stretches of time than used to be thought. Here as elsewhere we have learned that underlying forces work out their effects but slowly. As regards the phenomena which have been considered in the present Book, the thing most notable is the imperfection of competition within a country for equalizing the returns between industries of the domestic type and those of the international type. To put the same thing in another way, the output of goods and service which are solely or chiefly for domestic consumption does not easily or quickly come in such relation to the output of things which are largely exported that the returns in the two groups are even roughly the same. Even when those returns are for any reason made different to a serious extent, it takes a long, long time before there are shifts in the application of the productive factors great enough to bring about a stable situation. Hence it is that the dislocation of foreign exchanges may have more lasting effects than might be expected. And hence it is that the relative purchasing powers of the currencies (the purchasing-parity doctrine) has an effect like that of a deep undercurrent, obscured and impeded and deflected by upper movements which alone are easily observed, and which at any given time seem alone to have an effect on our well being or on the choice of a guiding policy.

CHAPTER 34

THE THEORY OF INTERNATIONAL TRADE. WHY PARTICULAR GOODS ARE EXPORTED OR IMPORTED. THE PRINCIPLE OF COMPARATIVE ADVANTAGE

§ 1. Some familiar facts: money incomes and prices differ in different countries; but prices of goods entering into international trade tend to be the same. Money wages not necessarily low in exporting countries.—§ 2. A country exports those things in which its labor is relatively effective—in which it has a comparative advantage. Illustrations from countries of high wages and of low wages.—§ 3. Differences in advantage may rest on climate or other physical cause, or on acquired aptitudes.—§ 4. Specially low wages of a particular class of laborers operate in the same way as does a comparative advantage in real cost. General low wages do not affect international trade or enable universal underselling.—§ 5. A country may import things for which its labor is productive, if its labor is even more productive for other things. But international trade rests largely on absolute differences.—§ 6. The gain from differences in comparative cost is dependent on immobility of labor between countries.—§ 7. A country may import part of the supply of a given commodity, produce a part at home. Difference between extractive and manufacturing industries in this regard.

§ 1. The preceding chapters have considered chiefly the mechanism of international trade. We may proceed now to matters more fundamental: the variations in prices and money incomes in different countries, the causes which determine what commodities a country shall import or export, the real gain from international trade. The first topics for consideration will be the mode in which a country's exports and imports are determined, and the cause and significance of variations in prices and incomes.

We shall assume, for simplicity, that trade is free. Duties on imports have important modifying effects, but these can be understood better if the working of unfettered trade is first examined. Further, we shall assume that all the trading countries are on the gold standard. The fact that they are not also has modifying effects,

which have been considered in the preceding chapter and do not affect the fundamentals.

Begin by calling to mind some familiar but often neglected facts, known to all observers but rightly interpreted by few. Among the most familiar is the existence of differences in the value of money in different countries; that is, differences in the range of prices and of money incomes. It will appear later in our inquiry that the differences in money incomes are the more important, and that prices do not always move with money incomes; but for the present we may assume that prices and money incomes in general move together. Money wages and other money incomes, and most prices also, are higher in the United States than in England, higher in England than in France and Germany, higher in these latter countries than in Italy and Spain; and lowest in countries like Japan, India, China.

While these differences in money incomes and in the prices of many goods are familiar, it is clear also that some commodities differ but little in price in the various countries. These are the commodities which are dealt with in international trade and move freely from country to country as imports and exports. They are the same in price in all the trading countries if we set aside cost of transportation. Where cost of transportation is considerable, their prices may vary considerably; hence it can only be said that the prices tend to be not far from the same. Some further qualifications to the general proposition would have to be made if it were attempted to fit it with exactness to all the facts. Sometimes an unfamiliar commodity goes from one country to another, is bought very cheap by traders in the one and sold very dear in the other; there is then a marked difference in price. This is likely to happen where industrial developments or rapid improvements in communications cause new opportunities for trade to arise. But in time mercantile competition tends to wipe out these differences. Here as in other directions the pioneers make money; unusual profits are presently cut away; in the end only such differences in prices persist as are accounted for by cost of transportation and the ordinary business returns. Goods which are the subjects of a constant and considerable foreign trade are sold at about the same

prices in all the trading countries. Wheat sells at approximately the same price in the United States and England, tea in the United States and Ceylon, coffee in the United States and Brazil, wool in Australia on the one hand, in England, France, Germany on the other. Nearly the same prices, be it noted. In order that a commodity shall move from one country to another, it must always be somewhat cheaper in the exporting country—cheaper at least by cost of transportation.

Money wages, however, are not necessarily lower in the exporting country. Thus they are higher in the United States than in England; yet the United States exports wheat to England. They are higher in England than in China; yet England sends all sorts of manufactured goods to China. They are higher in Australia than in Germany, yet Australia sends wool to Germany. A common notion in regard to international trade is that a country where wages are low is peculiarly likely to have large exports, and that one with high wages has difficulty in sending out its exports. Yet a moment's consideration of familiar facts such as have just been adduced shows that this cannot be the case. And the preceding chapters have shown that the exports of a country balance in money value its imports (barring those differences one way or the other which are explained by payments other than for merchandise). The countries with high money wages have no less exports than those with low money wages. In the trade between the two sets of countries neither can export more than the other; the payments between them balance.

§ 2. These preliminary matters point to the first important proposition with regard to international trade. A country exports the things which are low in price in its borders, and these are things in which its labor is applied effectively. Put in words more often used in the literature of economics, a country tends to export those things in which it has a comparative advantage. And conversely, a country tends to import those things which, if produced within its borders, would be high in price—those in which its labor would be applied with less effect, those in which it has a comparative disadvantage.

For example, wheat is exported steadily in large quantities from

the United States.[1] The money incomes of the farmers and their hired laborers who produce it are relatively high. If the price of wheat is low in the United States, it must be because the labor of those who produce it is effective. That is, the labor is applied to advantage. The common notion that high wages and high prices necessarily go together is quite unfounded; high wages are found with low prices if the productiveness of labor is great. Money wages are here referred to; as regards real wages, it will appear more fully as we go on that a high rate is the result of a widespread effectiveness or productiveness. Looking at money wages alone and considering them in relation to international trade, a high rate, if accompanied by great effectiveness, is no obstacle to low prices of a particular product and to its exportation. If, on the other hand, there be high money wages without any special effectiveness, then there will be high prices. The employer who must pay high money wages and whose laborers do not produce abundantly must sell his product at a high price in order to meet his expenses. In a country of high money wages the producers can continue to export in those branches of industry only in which the effectiveness of labor is great. The producers in those branches where the effectiveness is smaller will find more and more difficulty in meeting foreign competition, and may be driven out of business by competing foreign imports.

Again: China exports tea and raw silk; British India exports jute; Brazil, coffee. These are countries in which money wages are low. But they are also countries in which labor in general is ineffective. They import large quantities of manufactured goods which are produced more cheaply by effective and highly paid labor in the manufacturing countries. They export those things in which their labor is perhaps ineffective, but is *less* ineffective than it would be in making textiles, hardware, and other manufactures. They export those things in the making of which they

[1] I use wheat for illustration, tho the exports may cease in the course of time. The tendency to decline in these once heavy exports is commonly ascribed to the fact that we "need" the domestic product for our rapidly increasing population. This is true, as far as it goes. But the reason why the product fails to keep its former relation to population and "need" is the increasing cost (marginal cost) of wheat. That increase in cost means lessening effectiveness and lessening comparative advantage; hence lessening exports.

have a comparative advantage; that is, those for which, in their own borders, labor is most effective.

Thus we reach, alike for countries with high money incomes and with low money incomes, the same conclusion: those things are comparatively cheap, and those things are likely to figure in the exports, in which the country's labor is the more effective. This is the principle of comparative cost or comparative advantage which bulks so large in the literature on international trade; often misunderstood and its meaning and implications often neglected even tho it be at the outset understood and accepted. It is the theme to be developed with some elaboration in the rest of this chapter.

§ 3. It matters not, for the purpose in hand, what are the causes of the special effectiveness of labor which constitutes a country's advantage. It may arise from climatic superiority or other natural fitness, or from skill and aptitude acquired by man; or it may arise from a combination of these. The advantage of the United States in wheat, and its exports of wheat, rests (or rested) partly on the possession of vast tracts of new and fertile land; but it was much promoted also by the intelligence of its farmers and their large use of agricultural machinery, and by cheap rail transportation from the western wheat fields to the seaports. All sorts of causes here concur; not only the obviously natural ones, but those connected with land tenure and land ownership, with universal education and universal ambition, with the influence on freight rates of unfettered enterprise in the construction and management of railways. However complex the causes, their outcome is clear: the labor of producing and shipping American wheat is effective. The same complexity of causes lies back of our exports of petroleum and of copper—not merely great natural resources, but also great enterprise and skill in developing them. In some of our exporting industries, enterprise and skill alone, without unusual resources, suffice to explain effectiveness and cheapness; as in agricultural implements, sewing machines, automobiles, builders' hardware, electrical equipment. England's large exports of manufactures also illustrate the same combination of high money wages, effectiveness of labor, com-

parative advantage. The exports rest partly on her deposits of coal and iron ore—the natural foundations of manufactures—partly perhaps on a favoring climate, very largely on the vigor, enterprise and skill bred by free industry and free political institutions. China's advantage (or less disadvantage) in tea and raw silk is the result partly of climate, partly of skill and experience transmitted from generations to generations of patient workers. That this latter cause of advantage may be precarious is shown by the extent to which, in recent times, some rival countries have deprived China of her former position as almost the sole exporter of these articles. Ceylon has developed large exports of tea, Japan of raw silk, by systematic attention to the best ways of making labor effective in producing these things.

§ 4. Cheapness in price being the proximate element in determining exports, any cause or set of causes which makes a commodity cheap acts as an advantage and so promotes exports. True, the prevalence of a low general rate of wages neither promotes nor checks exports. But if one particular kind or grade of labor can be had at *exceptionally* low rates of pay, the commodity made by it is likely to be exported. Tho the labor be not of high effectiveness, the exceptional low wages constitute low expenses of production, low cost from the business man's point of view; and they conduce to low prices of the things made and are equivalent to a comparative advantage in real costs.

Interesting questions upon this aspect of the problem are presented by the exports of cotton from the United States. Climate, in its effects on the quality of the fiber, may go far to explain these exports. Social conditions also have been supposed to be an important factor. Before the Civil War, slavery was thought by many to explain the cotton trade of the South; it gave the advantage of very cheap labor. But the great growth of the exports since that war (when once the first years of turbulent transition had passed) shows that slavery in itself was not the one controlling cause. It remains true, however, that cotton continues to be grown mainly by negro labor, and that this is cheaper than most American labor. The maintenance of the exports may thus be ascribed in part to the persistence of social conditions derived from slavery.

On the other hand, this very negro labor, cheap tho it seems according to American standards, commands higher money wages than are current in Egypt, India, and other countries from which a competing supply of cotton comes to the world's markets. The labor must be at least to some degree effective. Further, much American cotton is grown (in Texas, for example) by white labor which earns the normally high American rates. Here the main explanation of the exports must be found in the effectiveness of the labor—climatic causes being no doubt important in contributing to that effectiveness.

Clearer illustration of the influence of specially low wages can be found elsewhere. In Saxony and Bavaria there are districts where a congested population was and apparently still is willing to work long hours for low wages. Toys and some sorts of textiles and knit goods are turned out at very low prices and are exported in considerable quantities. In England, again, while most exports rest on effectiveness with high wages, there are so-called "parasitic" industries (lace making and chain making are examples) in which wages are especially low and in which prices are low in consequence. From the social point of view, these are not welcome elements in a country's trade, whether domestic or foreign. But so far as the currents of international trade are concerned, specially low wages and specially productive labor operate in the same direction —both promote the exports of the commodities affected.

In these cases of low wages, the effect on international trade depends, to repeat, on their being limited cases. If *all* wages in a country are low, no one commodity will be cheaper than any other, and no effect on exports or imports will ensue. On this subject there is a sort of terror among many persons in countries of high wages like the United States and England—a fear of universal underselling and wholesale collapse because wages are lower in some countries from which exports come. The relation between international trade and general low wages is so commonly misunderstood as to deserve more than cursory consideration. The essentials are most easily brought out, here as elsewhere, by analyzing the simplest situation, trade between two countries having the same monetary standard.

Suppose two countries—say the United States and Japan—suddenly to open commercial relations, there having been previously no trade between them. Suppose money wages to be lower in all occupations in Japan, and all goods to be cheaper there. Money then has a higher value in that country than in the United States; trade in merchandise moves one way only, goods being sent to the United States; specie alone flows to Japan. Prices and wages will then rise in Japan, and will fall in the United States. As this transition goes on the flow of specie to the country of falling prices will gradually diminish, and will finally cease when equilibrium has been established. But that equilibrium will not necessarily be reached at a stage of equal wages in both countries; still less at a stage of equal prices in both, and consequent cessation of all trade between them. As prices in general move up in Japan, in response to the inflow of specie, it will appear that the prices of certain commodities do not move up to the American prices of the same commodities. These are the commodities in which Japanese labor is effective, or (possibly) in which some sets of Japanese laborers get unusually low wages. Such commodities will continue to be exported from Japan even after wages and prices in general have risen. Conversely, in the United States wages and prices will fall. But as they fall, some things will prove to fall in price below the new Japanese level. These are the things in which American labor has an advantage or (possibly) in which it must submit to specially low wages. Such things will begin to be exported to Japan as prices there rise, and they will continue to be steadily exported. The upshot is that there can hardly be such a thing as continued underselling in *all* goods. There will almost certainly be some approach toward equalization, perhaps complete equalization, of the value of money in the two countries; and thereafter a maintenance of imports and exports, each country exporting those things in which it has an advantage and importing those in which it has a disadvantage.

No such simple case has ever appeared. Economic events do not proceed as if in a laboratory where experiments are neatly arranged. The adjustment of relative wages and prices in different countries has taken place by a gradual and almost insensible

process. Possibly something like an abrupt change did take place in Japan in the last decades of the nineteenth century, when that country, previously sealed to foreigners, was opened to trade with them and entered on her amazing political and industrial transformation. In most cases, the main lines of adjustment were settled long ago. And this general adjustment, it should be noted, has by no means been such as to bring about an equalization of money incomes or of general prices; it has not brought about a uniform value of money the world over. In the supposed trade between the United States and Japan, equilibrium and settled exchange would doubtless be reached—the industrial characteristics of the two countries being as they now are—while money incomes and most prices were still higher in the United States. So it is in the world at large; differences in money wages and in the prices of domestic goods persist between country and country. What are the causes of these variations we have yet to consider. It is certain that they do not lead to universal underselling or to a continued trade in which goods move one way only.

§ 5. It follows from the principle of comparative advantage that a country may fail to produce things which at first sight one would expect it to produce with benefit. It may import things in which its labor is more effective than is labor in the country whence they come. Not all international trade rests on this precise relation; but under it the characteristics of international trade appear most markedly.

If a country, tho under no disadvantage in producing a commodity, has herein a *less* advantage than in others, that commodity will be imported. Labor in the United States is no less productive in growing hemp than labor in Russia; it is probably more so; none the less, hemp is imported from that country. Labor in the United States is no less effective in producing flax fiber than labor in Belgium, or in making linens than labor in Germany or Ireland; but flax and linen are still imported, and this in face of a considerable duty. Coarse wool, such as is used in making carpets, could be grown here with as little labor as in China, Asia Minor, Russia and sundry other backward countries, from which none the less it is steadily imported. The everyday ex-

planation of all these phenomena is that labor is too dear in the United States. The explanation is true enough, as far as it goes—but why is the labor dear? Our high rate of wages does not lead to the importation of all goods whatever, or prevent the exportation of those in which the productivity of labor is large. High general wages are the results of high general productivity. Once established and current, they constitute a difficulty for other possible industries in which productivity is not high. The real explanation of the continued importation of things in which labor is at no disadvantage is that they cannot meet the pace set by those in which the labor of the country is *more* productively applied.

Obviously, it is to the interest of a country to turn its labor into the most advantageous channels; not merely to those industries in which it is under no disadvantage or has only a slight advantage but to those in which it has the greatest advantage. Similarly an individual finds it to his advantage to devote himself to that occupation in which he is most proficient. The bricklayer does not carry his own bricks, even tho he could carry as many as the hodcarrier, perhaps more. He can lay the bricks better than the hodcarrier and gains by confining himself to that. An able business man delegates to clerks and subordinates much routine work, even work involving some responsibility and judgment, which he could do better himself; he confines himself to the still more difficult tasks of management in which he has peculiar excellence.

By no means all trade between countries, or all division of labor between individuals, is explicable in just this way. Often there is an absolute advantage on both sides. The bricklayer may be skillful without being physically strong; the hodcarrier may be able to carry more bricks. Each can do his own work better than the other. The capable business man may not be able to do clerical work as well as his bookkeeper. A certain impatience and abruptness of temper, characteristic of commanding personalities, may unfit him for monotonous office work. Similarly, a country may be at an absolute disadvantage in one industry and may have an absolute advantage in another. Such is the relative situation of temperate and of tropical countries as regards the articles they com-

monly exchange with each other. Brazil produces coffee with absolutely less labor than the United States could, India jute, and so on; while the United States produces wheat and makes it into wheat flour, spins and weaves cotton cloth, with less labor than they can.

§ 6. We must distinguish between the gain from international trade and in its probable extent, according as it rests on differences merely in comparative advantage, or on absolute differences of the sort last described. In the second case, where each country has a clear superiority, exchange between them will be to mutual advantage under any circumstances. Tho they may not share equally in the gain (of this more will be said in the next chapter), it will be to their interest to carry on the trade. But where there is a comparative advantage only, the existence of the trade, and the gain from it, rest on the fact that labor does not move freely from country to country. Suppose, for example, that labor in the United States is more productive all around than labor in Italy; it will be none the less to the advantage of Americans to import from Italy those things in which, tho they have an advantage, the advantage is less. But it might also be to the interest of the people of Italy to move *en masse* to the United States. Only because the Italians prefer to remain in their own country will that trade be carried on which does not rest on absolute advantage.

The indefinite continuance of trade explicable only from comparative advantage rests on immobility of labor between countries—on the ties of language, nationality, religion, on the obstacles from ignorance and poverty, which hold people to the land of their birth. Great as is the emigration of modern times, it has not sufficed to put an end to this prevailing immobility. In the example just given, if all the Italians were to move to the United States, they would be better off than before. So long as they remain at home, they provide the Americans with goods more cheaply than these could be made in the United States. Once in the United States, they would indeed produce the commodities with less labor than before; but that labor would have to be paid for at the higher American rate, and the commodities would be less cheap. The Americans (let us say, in courtesy, the other Americans) would

be less well off. It is conceivable, to be sure, that when the Italians got to the United States, they would not receive the full American rate of wages. They might continue to work for the Americans, as they had done in Italy, at a low rate of wages. And it is true that our newly arrived immigrants, in fact, are in a group by themselves. But their pay shows at least some approach to the American rate. Tho they work for wages not up to the level normal in the United States, they do not work for us as cheaply as do their countrymen who remain at home. It happens, also, that often they do not work at the identical things which are imported (or under free trade would be imported) from the old countries. These things it continues to be to our advantage to procure by the way of foreign trade, tho our labor may be more effective in making them than is the foreign labor. In an ideal—we might say utopian—distribution of the world's productive forces, the division of labor and trade which rests solely on comparative differences in costs would not exist. But as men and nations are, no small part is played by the great historical gulfs between nations and races and by the resulting immobility of labor.[1]

§ 7. From the preceding exposition, it might be inferred that a country produces within its own borders no articles which it imports, and that, conversely, whatever articles it exports are supplied *in toto* to the other country or countries. But this does not necessarily follow. More especially it does not follow with regard to the considerable range of commodities which are produced under the conditions of varying costs or diminishing returns.

Take the case of wheat, which the United States and Canada export to England and Germany. Some wheat can be grown to advantage in these countries—a great deal in Germany, less in humid England. They are at a disadvantage only when they force the margin of cultivation and raise wheat on the land less adapted for it. On the better sources of supply, the domestic producers can hold their own, even tho imports come in. Hence the national supply comes partly by importation, partly by domestic production. The same is the case in the United States with wool. Some

[1] This topic connects itself with the general subject of differences of wages and non-competing groups, and with some things fundamental in the theory of value which have not yet been touched; see below, Chapters 52 and 53, in Vol. II.

parts of the country have clear advantages for wool growing and are adapted for little else—such as the semi-arid plains of New Mexico and Arizona. Elsewhere the climate is not so favorable, or (what is more important) the land can more advantageously be put to other uses. Corn and wheat pay better; there is merely a comparative disadvantage in wool growing. The total supply of wool which the country wants at current prices cannot be produced in those regions which are advantageous enough; hence part is imported. This sort of importation—wool into the United States, wheat into Germany—takes place notwithstanding duties of considerable weight on these products in the two countries. As the better sources of supply in each country have been fully utilized, it has become more costly to procure additional wheat and wool; hence, as the poorer sources are resorted to, the price rises until imports come in over the duty. Most of the supply in each country comes from domestic production; but there is a steady importation.

With manufactured goods the case is somewhat different, since there are not commonly the same limitations set by nature to the increase of supply at constant cost. It is true that some division of the field is likely to take place here also. Not infrequently permanent causes of variation of cost exist between different establishments,—in the iron manufacture, for example, as regards supplies of ore and coal. And even when there are not such deeply rooted causes of variation, forces of a similar sort are in operation for considerable periods. The principle of constant returns, tho it works out its results for most manufactures in the long run, is subject to great modification in times of rapid change, such as the modern world has seen in so many cases. At any one time, some establishments in a given industry, say the woolen manufacture, may have such advantages as will enable them to hold their own against foreign competition, and others may not. Part of the supply, but not all of it, will be got by importation.

Nevertheless, division of the field between foreign and domestic manufacturers is less common and less likely to persist than such division between foreign and domestic producers in extractive industries. True, it happens occasionally that there is a very even

balance between the two, and that minor factors, such as established name and repute, or skill in satisfying the tastes and whims of consumers, will determine that some among each set of producers will hold their own in the market. More often, when part of the supply of a manufactured article, and part only, is recorded as coming from abroad, it will be found that the imported article, tho it has the same name as the domestic, is of such a different quality as to be virtually a different thing. French woolens are imported into England; but they are different goods from those which the English make and export. The same explanation of an apparently anomalous phenomenon (simultaneous import and export of the same article) applies to some raw materials. Tho the United States is a great exporter of cotton, she imports cotton also; but it is not the same as is exported. What comes in is Egyptian cotton, of longer fiber than the ordinary domestic, used for certain knit goods and other fabrics.

CHAPTER 35

THE THEORY OF INTERNATIONAL TRADE (*Continued*). WHEREIN THE GAIN CONSISTS

§ 1. Difference between trade within a country and international trade. Varying rates of wages in different countries show varying gain from the exchanges between them.—§ 2. An illustrative case, England and Italy. Demand and utility determine relative wages and prices. Slow and obscured operation of this cause, thru the influence of the money supply on incomes and prices.—§ 3. Effects of changes in international demand; of new articles of export; of payments other than for merchandise.—§ 4. Difficulty of following these causes in detail, illustrated by the case of the United States since 1873.—§ 5. Money incomes, not prices, important in determining the gain from international trade.—§ 6. Two causes act on the gain: the play of international demand and the effectiveness of labor in producing exported goods. The last cause settles the general rate of money wages.—§ 7. High money wages and other incomes do not necessarily bring high domestic prices. Illustration from the United States.

§ 1. In the preceding chapter the general nature of the gain from international trade, and the causes that determine which among a country's commodities will be exported, have been considered. It remains to examine more carefully the nature of the gain and the manner in which it is shared by the trading countries.

For this phase of the question, it will be best to turn first to the case where each country has an absolute advantage in the commodities it exports. Such is the nature of most trade between tropical and temperate countries. Such also is the nature of most trade between different parts of the same country.

Within any one country this sort of geographical division of labor does not commonly give rise to any peculiar problems. Exchange takes place between the various sections, but as a rule on equal terms. Within a country, there are no very great variations in wages and incomes—for persons of the same capacity and skill—between different districts. Yorkshire is a great woolen manufacturing region, Lancashire is a great cotton goods district; they exchange products; but wages are substantially the same in the two, and they share equally in the advantages of the exchange.

Pennsylvania produces iron and coal, the Mississippi Valley wheat and corn, Oregon lumber, California fruit, the South cotton, New England sundry manufactures. Tho there is no complete equalization of incomes between different parts of the United States, we find homogeneity in the country's general industrial conditions. Labor flows with much freedom from one part to the other (unless, indeed, it be kept from parts of the South by the race complication), and there is a generally high level of money wages. Not only is the geographical division of labor determined in the main by obvious advantages in production, but the people of all parts share to about the same degree in the general cheapness and abundance which it promotes.

But between tropical and temperate countries, and between countries civilized and those half civilized, there is no approach to equalization of incomes. India, China, Japan, South America have very different rates of wages from the United States and Europe. And there are also differences between the United States and European countries, and between the various European countries. Money incomes being different, the gain from international trade is unequally divided. The commodities exchanged—the international commodities—sell at the same price (barring cost of transportation) in the several trading countries. The English, for example, buy woolen goods and cotton goods and coal, as well as fruits and wines, as cheaply as the Italians. But the English have higher money incomes with which to purchase both sets of commodities, and hence benefit more from the trade between the two countries than the Italians do. How explain these differences?

§ 2. Suppose an extreme and simple case, typified by England with exports of coal, Italy with exports of lemons; these being exports which result from obvious advantages in production possessed by the two countries. The quantity of coal which England will send to Italy depends on the conditions of demand in that country. In Italy the coal will sell at the English price, plus cost of transportation, and at this price a certain amount of coal can be disposed of there. In England, on the other hand, lemons will sell at the Italian price plus cost of transportation, and at this price a

certain quantity of lemons can be disposed of in England. If this exchange has been steadily maintained during a long period, it is probable that the two amounts will just balance—the coal exports from England will just equal in money value the lemon imports into England. But if we suppose the two countries to be suddenly brought to trade with each other, no previous process of adjustment having taken place, this would be highly improbable. One or the other sum is likely to be the greater. Suppose the lemons exceed. Then Italy will export to England more (in money value) than she imports. Specie will flow to Italy. Prices will rise there, and money incomes will rise with them. Prices and money incomes, on the other hand, will fall in England. As this process takes place, the export of lemons from Italy will be checked; for lemons rise in price there with other things, while the English consumers who buy them have lessened money incomes. But exports of coal from England will be stimulated. Prices are falling in that country, and the price of coal falls with other things; hence coal can be sent to Italy at a lower price. Its consumption in Italy is stimulated, not only by its lower price, but by the fact that money incomes in Italy are rising. Eventually a stage is reached at which the coal just pays for the lemons. The imports equal the exports in money value, specie no longer flows, equilibrium is established.

Just at what point this stage is reached evidently depends on the nature of the demand for the two articles in the trading countries. If the demand for both is elastic, equilibrium will be quickly reached. If the higher price of lemons quickly checks the English consumption, and if the lower price of coal quickly stimulates the Italian consumption, the money values of the two articles will soon become equal. But if the people of England have a strong demand—that is, an inelastic demand—for lemons, and continue to use very nearly the same quantity of them, even tho their price rises somewhat; and if, on the other hand, the people of Italy have an inelastic demand for coal, and do not increase their use of it, even tho its price becomes lower—then a long process of changing prices and changing incomes will ensue. The country whose demand is great for the products of the other coun-

try will have comparatively low prices and low money incomes; the country whose exports are in insistent demand in the other country will have comparatively high money wages and money incomes.

The case supposed (imaginary in that it assumes these two articles to be the only ones exchanged between England and Italy) points to the underlying principle. That country gains most from international trade whose exports are most in demand, and which itself has little demand for the things which it imports—that is, for the exports of other countries. That country gains least which has the most insistent demand for the products of other countries. In the semi-technical terms which we have used elsewhere, we may say that the extent to which a country shares in the gain from international trade depends on the marginal utility of imported goods to its people and the marginal utility of its exported goods to foreigners.

This cause operates thru the mechanism for the distribution of the world's stocks of specie and the equalization of international payments. It need not again be said how slowly this mechanism moves, how difficult it is to follow its operation in the confused currents of international trade. The imports and exports of the various nations have long accommodated themselves to a scheme (if that word can be applied to something which has grown up without plan or intent) of differences in the value of money in the various countries. Tho changes in the scheme take place, they come by slow and half-concealed movements. Many economists of modern times, intent only on those phenomena which are unmistakable and susceptible of exact observation, ignore the less conspicuous underlying forces, and are skeptical concerning the validity of fine-spun theories about them. Yet the broad phenomena are explicable only by reasoning of this kind. As was pointed out at the very beginning of the discussion of international trade, it is obvious that great differences in money incomes exist between the various countries, and that they persist thru very long periods. It is obvious, too, that these differences bring inequalities in the gains from foreign trade. The flow of specie, again, is governed by the prices and the consumption of the articles that enter into

foreign trade. Notwithstanding that flow, the value of money is not brought to the same level the world over; and notwithstanding the variations in the money incomes and prices, substantial equilibrium in payments between the countries is still reached. All these phenomena are brought into orderly connection by the theory of reciprocal demand.

§ 3. Suppose now another case, nearer to reality. Suppose that, between countries whose trade has previously come to equilibrium, a change in demand sets in; that England, for example, having imported from the United States as much as her exports paid for, now demands more of American goods. Stated in more specific terms, the supposition is that English consumers buy at the ruling prices more of the American goods—cotton, wheat, copper, or what not—than they bought before. Specie then must flow from England to the United States; or, what amounts in effect to the same thing, new specie from the mines which would otherwise have gone to England must be diverted to the United States. If this movement takes place on a considerable scale and for a considerable time, it must affect prices. The same train of consequences will ensue as in the supposed trade between England and Italy. Prices and money incomes will rise in the United States, and will fall in England. Eventually these shifts will again bring equilibrium. The higher American prices will gradually check the increase of exports from the United States, the lower English prices will stimulate an increase of exports from England. Money incomes will reach a somewhat higher level in the United States, a somewhat lower level in England. As consumers of English goods, the Americans will gain; as consumers of American goods, the English will lose. Thus the increase in English demand for American goods will cause the English to gain less from the trade between the countries, the Americans to gain more.

Again, the appearance of a new article of export in a country's foreign trade operates in the same way. Kerosene oil became an important article of export from the United States after the Civil War—a clear addition to the things which foreign consumers bought. This had to be paid for. If trade before was at equilibrium, and if no other disturbing factor entered, specie must have been

diverted to the United States, as in the case previously supposed. The same consequences must have followed until, by a gradual stimulus to foreign imports into the United States and by a gradual check to exports (other than the new article) from the United States, equilibrium was reëstablished, with a new scale of money wages and domestic prices in the two countries. The foreign countries indeed gained by having the new article which they did not have before. But they lost by having to pay somewhat higher prices for other American articles, and by having somewhat lower money incomes with which to pay for them.

An obligation to make other payments than those for merchandise has corresponding effects. If a country has remittances to make to other countries—whether for travelers' expenses, interest on accumulated debt, freight charges, or like debit items —it is likely to be in a worse position as regards the gain from its trade with the other countries. The remittances must be made in money, or in money's worth. They must be made in the first instance—international payments having previously balanced—by an outflow of specie. That outflow of specie lowers prices; it stimulates exports and checks imports. In the end the payments are effected by an excess of merchandise exports. But the process which brings these added exports brings also lower prices and lower money incomes in the remitting country, and so lessened advantage from international trade. The need of forcing more exports on the foreign consumers causes the foreigners to get the exports on better terms, and causes domestic consumers to get the foreign imports on worse terms.

§ 4. It is extremely difficult to follow these forces in any concrete case; for it is rare that any one factor operates alone, even rare that several factors combine to operate in the same direction. Nowhere is this difficulty better illustrated than in the experience of the United States during the last quarter of the nineteenth century. The play of international demand works out its results over long periods—it is only thus that the flow of specie affects prices. But during the period mentioned, a multitude of conflicting forces have been at work. After 1873 the foreign trade of the United States in one respect clearly took a new turn: the

merchandise exports, which previously were less than the imports, came to exceed the imports.[1] The change is explained by the various additional payments (for interest, travelers' expenses, immigrants' remittances, freight charges, and so on) which Americans had to make to foreigners. This circumstance would tend to cause an outflow of specie, to lower prices and incomes, to make the terms of international exchange less favorable. But during the same period (after 1873) a great increase in the demand for American exports set in,—for cotton, wheat, meat products, and for some manufactures; while new articles of export, such as kerosene and copper, became important. All this worked in just the contrary direction. During this period, moreover, a policy of protective import duties was applied with great rigor; and such a policy also, as will appear presently, promotes a higher scale of money wages and prices. Meanwhile gold was steadily mined on a large scale within the country. Legislation also was an important factor in the country's monetary supply: consider the resumption of specie payments, the injection of silver money from the acts of 1878 and 1890, the peculiar working of the national bank system. Thruout the period there was great growth of population and wealth, and therefore a great increase in the demand for money—an increase which, taken by itself, would tend to lower prices. What a jumble of interacting and conflicting elements! How say whether the forces that made for greater gain from foreign trade outweighed those that made for less gain? How follow in detail the concrete working of any one factor? The difficulty would probably be the same in kind, tho less in degree, if an examination were attempted of the foreign trade over a considerable period of any European country.

The difficulty is one common in economics. A number of forces —acting sometimes in the same direction, sometimes in conflict with each other—combine to bring about a given result. Being necessarily debarred from deliberate experiment, we must resort to hypothetical reasoning, and must be content with general conclusions confirmed only in part by direct experience. Thus, we reason that an increase in the money supply must raise prices.

[1] See above, Chapter 32, § 1.

We find that in the long run, and making due allowance for credit fluctuations, for bank reserves and bank expansion, this is true; and the result is further confirmed by the striking phenomena of paper money inflation. We reason that a flow of specie from one country to another tends automatically to bring its own check, and that payments between countries tend to balance without the movement of specie. We find that in fact payments are usually adjusted with a very small use of specie; while those cases in which it moves steadily one way—from specie-mining countries, for example, or from western countries to the sluggish communities of the East—are exceptions of the sort that confirm the rule; they are explicable on grounds of their own. We reason that the stage of equilibrium in payments is reached by a process which involves in the several countries different levels of money incomes and prices; and we find that in fact some countries have markedly higher wages and prices than others. All these verifications of the general reasoning give us confidence in phases of the reasoning which we cannot verify directly. Among the conclusions verified in this indirect way is that reached in the preceding discussion: a country's share in the gains from international trade depends on the play of reciprocal demand. The more insistent is the demand for a country's products in other countries, and the less insistent is its own demand for the products of other countries, the greater then is its gain from international trade.

§ 5. The rise and fall of money incomes and of prices, in relation to international trade, call for some further discussion. Strictly speaking, it is not the rise or the fall of prices, but that of money incomes alone which is of consequence.

Differences in the value of money—in the general level of prices and incomes—which result from the play of international demand are of real and permanent importance only with reference to foreign goods. A rise of prices and incomes all around is in the long run immaterial; it means only the use of more counters in exchange. It is true that the process of transition disturbs the relations of debtors and creditors; perhaps true also that rising prices bring a certain stimulus to production. But these are transitory effects. In the end people are no better off from having higher

money incomes if prices rise to the same extent. Now when changed conditions of international trade bring about higher prices, domestic prices will rise as well as wages and other incomes; nothing is gained in substance as regards the domestic goods. But prices of foreign (imported) goods are differently affected by these same conditions. They do not rise; they fall. The higher incomes go further in the purchase of foreign goods, and in these purchases only. Conversely, a fall in incomes and prices, the result of changes in international trade working in the opposite direction, affects consumers only in their purchases of imported things. As regards domestic commodities, the fall in money incomes is offset by the fall in their prices. But foreign goods tend to become dearer, and in buying these there is a real lessening of the gain from international trade.

Changes in the value of money caused by varying currents of international trade thus bring about not only transitional effects (such as those on debtors and creditors) but permanent effects as well. But these permanent effects are of a different sort from what is implied by the phrases commonly used. They do not arise for example from the fact that increased exports bring in more money. They arise because foreign goods are secured on easier terms. An increase in the monetary supply, equally distributed over all the world, would benefit no one. But an increase which went to one country alone, or which went in larger proportion to one country than to others, would benefit the people of that country in their dealings with other peoples. And similarly a redistribution of the existing supplies, caused by changed conditions of demand in the trade between the different countries, would cause the people of some to secure greater gains from their dealings with others.

One application of this reasoning is obvious. A country gets its gain from international trade only if it takes advantage of the relative cheapness of foreign goods. So long as these are admitted free of duty, and only so long, does it secure in full the real advantage from high money incomes or from a rise in money incomes. Duties imposed on foreign goods simply cut off that advantage. And if the duties operate to bring about the production within the country of goods which, but for the duties, would be

imported, the gain from foreign trade entirely disappears. Such, in brief, is the main argument in favor of free trade; and, as far as it goes, it is unanswerable. It is by no means all that is to be said on the controversy between free traders and protectionists, but it is a fundamental truth, much befogged by current fallacies, yet not to be mistaken by any one who has grasped the principles of division of labor, exchange, money, and prices.

§ 6. The extent of a country's gain from international trade depends on two causes: first, the terms of international exchange as just explained; second, the effectiveness of its labor in producing exported commodities. Both causes contribute to bring about the range of money incomes, high or low as the case may be, and so a greater or a less gain from the purchase of foreign commodities.

The action of the two causes is illustrated by the different positions of the United States and Russia as exporters of wheat.[1] The wheat growers in both countries sell their product at the same price in the world's markets. So far as the play of international demand goes, the people of both gain to the same extent. But so far as the cost of producing wheat goes—that is, the real cost, measured by the amount of labor needed for producing it—they are in very different positions. Wheat is grown with much less labor in the United States, and money wages are higher here. Money wages are lower in Russia, and the wheat growers of Russia, and indeed all Russians, are by so much in a less advantageous position in buying foreign goods. As between any two or more countries competing in the sale of the same article, the extent of their several gains from international trade depends on the relative effectiveness of their labor in producing the things exported.

The determining cause of the general rate of money incomes and wages in a country is to be found in the exporting industries. These set the pace; not for real wages but for money wages. Whatever is yielded by them tends to become, under the influence of competition, the ruling rate in the country at large—in other

[1] I believe (tho one cannot be sure) that the differences at the present time (1939) remain substantially of the same kind, and the illustration is still pertinent, notwithstanding the marked changes of agricultural conditions in both countries since 1909 when this passage was written.

industries as well as in those exporting. In the other (strictly domestic) industries, this money rate, to be sure, is a matter of comparative indifference, since the prices of commodities will rise and fall with the rise and fall of wages and incomes. The same parallel movement appears in the exporting industries—the prices of exported commodities go up and down with the money wages of the laborers engaged in producing them. And these money wages are *derived* from the prices at which the exported commodities are disposed of in the world's markets. No such parallel movement appears in the case of imported goods; lower prices of imported articles go with higher money incomes, and therein lies the real gain.

§ 7. A further question arises, in regard to which also there is much misconception. Is it always the case that a country of higher money incomes—that is, one with advantageous terms of international trade—is also a country of prices higher all around? Most persons would answer the question in the affirmative. But no unqualified answer can be given. It depends.

The key to the answer is in the distinction, now familiar to the reader, between the two classes of commodities: those which do enter into foreign trade, and those which do not. The former we may call, for brevity, international commodities; the latter, domestic commodities. Under conditions of free exchange, and with due allowance for the expense of transportation, international commodities tend to be at the same price the world over. Domestic commodities, however, may be at varying prices in different countries. And the range of domestic commodities remains wide notwithstanding the cheapening of transportation and the consequent extension of international trade and international competition. Many things are too bulky in proportion to their value to be moved far from the place of production; such are brick and stone. Some are so much affected by rooted habit that only the near-by producers can fashion them in the way desired by consumers; such are many articles of household furniture. Some are of necessity made on the spot where they are used; house accommodation is an obvious case. The services of physicians, lawyers, actors, musicians, domestic servants are also necessarily rendered

on the spot. These utilities (services) are of no small importance, especially for the well-to-do; their price evidently is determined by domestic conditions alone.

Consider now some domestic commodities, such as household furniture,—tables, chairs, bedsteads, chests. Will these be dearer in the United States, a country of high money incomes, than in Germany, a country of comparatively low money incomes? The answer depends on the effectiveness of American labor in producing them. If American labor is relatively as effective in this field as it is in foreign commodities, they will not be dearer. We have seen that American labor is more effective than German labor as regards wheat; otherwise, wheat could not be cheaper in the United States and could not be sent thence to Germany. But American labor may also be more effective than German as regards tables and chairs; and then tables and chairs, tho the laborers who make them get higher wages, will not be dearer in the United States. The principle is simple: those domestic commodities as to which a country's labor has the same degree of effectiveness as it has in making exported articles, will be relatively cheap, just as the exported commodities are relatively cheap. Those domestic commodities in which there is no such advantage will be dearer, and will be dearer to the degree in which the effectiveness of labor is less. The reader can supply for himself the extension of the argument which comes from the fact that some labor in a country, tho not effective, is paid at an unusually low rate. Domestic commodities made by such laborers will also be relatively cheap.

There is a common impression that the United States, a country of high money incomes, is also a country of high prices. But this impression rests on no certain basis. It has arisen probably because of the fact that many things are really dearer for the well-to-do. Services are almost necessarily dearer in a country of high incomes. Domestic servants, for example, get higher wages than in Europe. Physicians and lawyers get higher fees, teachers higher salaries. There are many things in which personal service, while not the sole element, is yet the most important; such are cab service and hotel accommodation. A great part of the income of the prosperous classes is spent on various forms of personal service, and

for these classes the "standard of living" (which means the expense of a given conventional mode of life) is high. Therefore those among the Americans or English who have fixed incomes find that their incomes go farther if they live abroad; hence their impression that *all* things are cheaper abroad. But many domestic commodities of general consumption are as cheap, or only a little dearer. Most food is equally cheap—not only that which enters into foreign trade, but that which is used solely at home. Fuel is as cheap in the greater part of the United States and in England as it is in Europe. As to the important item of house accommodation (house rent) it is not easy to make a comparison, because of the difficulty of making allowance for quality. I suspect that, taking into account size, convenience, and attractiveness, prices are not higher in most parts of the United States, for the housing accommodations of the masses; tho they doubtless are so for the rich, whose houses are built "by the day" and with less use of factory-made frames, doors, and windows. Clothing, and especially woolen clothing, is dearer in the United States—a result mainly of our policy of high import duties, which prevent us from using our high money incomes to advantage in the purchase of cheaper foreign woolens.[1]

[1] In this chapter and in that preceding, it has been tacitly assumed that within a country (so far as domestic commodities are concerned) exchange takes place, and value is determined, on the basis of labor costs—that value rests on "cost of production," not on "expenses of production." Elsewhere, however, it has been assumed that supply price, in its relation to value, means expenses of production, not "real" cost (see Chapter 12, § 1). The explanation of the inconsistency and the grounds for considering it not repugnant to the general validity of the reasoning upon international trade must be left for later discussion. See Chapters 52 and 53, Vol. II, for the further consideration of this subject.

CHAPTER 36

PROTECTION AND FREE TRADE. THE CASE FOR FREE TRADE

§ 1. The main argument for free trade is simple. Persistence of Mercantilist notions.—§ 2. Some popular arguments for protection: creating a home market; the truck farm case; creating employment.—§ 3. Protective duties as a means of combating industrial depression.—§ 4. The effect of protection on wages. Real wages in general lowered, tho some particular wages possibly kept up.—§ 5. The principle of equalizing cost of production.—§ 6. Effects of duties on prices and on consumers. A national loss only if domestic products are substituted for those imported. Monopoly may bring special gain to domestic capitalists, but brings no national loss. Labor monopoly may bring special gains to particular laborers.

§ 1. THE main argument in favor of free trade between nations has been already indicated. It is a simple corollary from the principles of the division of labor. Exchange between individuals brings the same kind of gain whether they live in the same village or in widely separated districts; things are obtained more easily and abundantly than by each person's producing for himself. The reasoning which shows that it is advantageous for the farmer to exchange with the village blacksmith, Maine with Florida, New England with the Mississippi Valley, California with the rest of the Union, makes out a strong *prima facie* case in favor of free exchange between the United States and England, between France and Germany. The burden of proof may be fairly said to rest on those who assert there is gain from the contrary policy.

Most of the common arguments in favor of restrictions upon trade, by protective duties or otherwise, are fallacious. Many are crudely Mercantilistic, resting on an assumption that imports are bad and exports good. The so-called unfavorable balance of trade is made much of; what is expended on imports is deemed so much wasted or lost. It is supposed that a decline in imports or an increase of exports necessarily brings money into the country; and the notion persists that herein there is a gain—one which results directly from a balance of money secured, not thru those recondite

effects on money incomes and foreign prices which were analyzed in the preceding chapter. Few among those who speak of a gain in exports as profitable ever heard of the last-named process or are able, unprepared, to understand it. They think of exports as bringing in money, and imports as taking money out, and money is the be-all and end-all of their economic thinking. Even if it is pointed out that a continuing excess of exports is due to other than merchandise transactions and does not bring in specie, the notion still persists that exports somehow mean gain and imports loss. The elementary truth that exports are but a means of procuring the imports on easier terms than the same goods could be got by making them at home—this is rarely grasped, or, if once grasped, is soon let slip.[1]

Mercantilist notions, universally discarded tho they are by the well informed, affect the policy of nations not only by strengthening the movement toward protection, but in other ways also. The public railways of most countries make special rates for exported goods, on the theory that this sort of movement deserves especially to be fostered. In the United States, the rate regulating authority —the Interstate Commerce Commission—sanctioned the same principle. Shipping subsidies are granted by many countries, and colonies acquired and maintained at great expense, with the same object in view. The United States government spends large sums in gathering information about opportunities for export and in promoting otherwise the export market; while various semi-public agencies and museums coöperate for this supposedly praiseworthy object. Underlying almost all activity of this sort is the persistent belief that there is something peculiarly profitable in international trade, and that the profit appears in the sale of the exports—a belief which exaggerates the importance of the trade and misconceives the nature of the real gain from it.

[1] What is true in this regard of the common talk of business men and newspaper writers is hardly less true of legislators and other public men. When I was in Paris in 1919, deputed to serve in the re-ordering of trade relations after the war, I found it quite hopeless to attempt to carry on discussion in other than the bald mercantilistic terms. All the representatives of the various allied countries presented their wishes and claims once for all on the basis that any measure promoting the exports of a country was *ipso facto* beneficial and any permitting larger imports *ipso facto* harmful.

Perhaps the ancient association of foreigner with enemy still lingers. People do not worry when New England buys coal from Pennsylvania; but when coal is bought from Nova Scotia, dire consequences are supposed to ensue. Half a century ago (more or less) the region which is now British Columbia was claimed by the United States to be part of its territory. Had the Oregon question been settled at that time in accord with the American claims, no one would have questioned that the resources of British Columbia in lumber, coal, and fisheries were of advantage to Americans. But once a border line is drawn, the situation is supposed to change; and that which would have brought us gain in the way of more abundant and cheaper supplies is fraught with peril precisely because these supplies came from a foreigner.

§ 2. Some of the popular arguments in favor of protection call for brief consideration: that it creates a home market; that it makes employment; and that it raises wages or keeps them high.

When imports are checked, and the things previously imported are made at home, a home market is supposed to be created. It is created; but is not, as protectionists commonly state or imply, an additional market. Another and different market is substituted. Here again most people's ideas do not get beyond the range of sales and of money dealings. When the linen manufacture (say) is established, those engaged in it buy food and other supplies; and here, it is supposed, is an additional market for food. The real "market"—that is, the real exchange—is of food for linens. That same market existed when linens were imported and food or other things were exported in payment. To cut off imports is to cut off exports also; it means simply the substitution of exchange within the country for exchange between countries. The essential question is whether for a given quantity of food (i.e. of labor exerted in producing that quantity) more linen is got in one way than in the other. The very fact that linen can be got cheaper by importation indicates that the foreign market is better than the domestic. The home market argument is most frequently used in the United States with reference to the farmers, who are supposed to get benefit from a greater demand for their products because of the establishment of manufactures. The presumption is, however, that they

do not gain but lose; the "market" which is created offers less in exchange for their products than does the foreign market.

A special form of the home market argument, also much used in the United States, is suggested by the truck farm. Suppose a manufacturing town is established in consequence of protection: the near-by farmers profit by the sale of milk, vegetables, and the like. These farmers do in fact profit; but simply because, while they sell all their produce in the town, they purchase from that town a very small share, if any, of the particular things there made. If they had previously exported all their vegetables and dairy products, and if the manufacturing town, after the duty, supplied precisely the goods which they had previously procured by importation, they would lose, not gain. The truck farmers, in truth, are ordinarily within the limited circle of real beneficiaries from protection. They gain, however, not as farmers, but as landowners. They are like the lucky holder of sites in or near to a newly established town. The great mass of farmers do not gain, but lose—those who supply most of the needs of the manufacturing population and who buy most of its products. The non-landholding people of the manufacturing town also fail to gain. As will appear more fully in the sequel, neither employers nor workmen are permanently better off. Only those gain in the end whose sites, whether agricultural or urban, happen to be more advantageously situated under the new distribution of the population.

Closely connected with the home market argument is that about employment. That protective duties add to the demand for labor seems patent to the everyday man and especially to the workingman. When imports are kept out, is it not clear that more employment exists for the workmen who make at home the things formerly imported? Here again people see only the first and most obvious results, and do not stop to think what other results must follow. If there are less imports, there will be less exports; and labor, if employed more in the new way, is employed less in the old. One of the most persistent of economic errors is the notion that employment is an end, not a means; and one of the hardest things to fasten in the average person's thinking is that the end to which employment should be directed is the increase of the national in-

come—the total flow of consumable goods and of services which constitutes the real revenue of the community. Most workingmen oppose labor-saving appliances and welcome arrangements which seem to increase the demand for labor. And most of them are instinctively protectionists, since the same fallacies are current in arguments for protection as in arguments for increasing the employment of labor. The workmen of any one group or set are concerned solely with their own share of the national income. Anything which adds, or seems to add, to the demand for their particular kind of labor is of course welcomed; and then, by an easy transition from the particular to the general, it is inferred that all labor is more in demand because of the circumstances which increase the demand in this particular direction.

One form of this argument—that employment is created—alleges that there is always unemployed labor and always unemployed capital. Put on a duty, bring this labor and capital together for making an article previously imported—and is there not a gain? The answer is that this problem is far removed from the protective controversy. Unemployed labor is a grave social evil; unemployed capital is a real waste. Some proportion of unemployment, no doubt, is inevitable both for labor and for capital; it results from shifts between occupations, from the processes of change and transition from progress in industry. To minimize it is among the most important of public tasks; it is also among the most difficult. But there is no ground for supposing that a system of protection would permanently affect it one way or the other.

If a new industry is stimulated in a country by a protective duty, it by no means follows that the labor which is unemployed is adapted to that particular industry or is in a place where it can take advantage of the new opportunities. Those who are out of a job cannot drop at once into the new places provided. Transfer and adaptation require time. And even supposing the improbable outcome that the unemployed labor and capital were really brought together in an industry created by protection—the solution of the problem would be but temporary. Inventions and improvements, redistribution of industries and of population, crises

with all their dislocating effects, would ere long cause the problem to present itself again. A country quite without international trade, shut within its own borders, would still be confronted with unemployment, as with other evils, so long as its industry rested on private property, complex division of labor, free movement of labor and capital; so long as there were hopes and fears, successes and failures, in the business world.

§ 3. The great depression of the third decade of this century led to a new turn in the use of the make-employment argument for the support of protective duties. The appalling numbers of the unemployed spurred endeavors to find work for them not only in the fields traditionally thought proper for public expenditure —roads, bridges, buildings for the civil service, men-of-war—but also in operations of the "private" kind such as workman's dwellings. In these operations the prime object was to create employment. While the resulting product was to be as abundant or useful as possible, the accruing utility was a secondary matter; the main thing was to provide work. Surely it was better, both for the unemployed and the rest of the community, that they should do something than that they should be supported in idleness. Did it not follow that one effective way was to impose duties on imports which would cause things to be made within the country rather than brought in from abroad? True, it might not be easy—as has just been indicated—to fit the unemployed labor into the new industries. Perhaps also, even after the labor had been placed and trained, the product might not be turned out as cheaply as it could be procured by importation. Still, was it not better to employ the labor in this way than to let it stand idle? and did not some net addition to the community's real income thereby accrue? The answer would seem to be quite the same for protective duties as for any of the other works promoted by public funds.

And yet the answer is not quite the same; there is a ground for distinction. The construction of roads, bridges, public buildings, even workmen's dwellings is not likely to continue indefinitely on a suddenly enlarged scale. Sooner or later people will see that this cannot go on without end,—that after all it is but a stop-gap. One may feel some assurance that it will be followed by a period of

quiescence; after a while operations of the kind, as they are resumed, will proceed on the scale adapted to current needs. It becomes plain that the sudden huge expenditure is "wasteful," that is, entails a comparatively disadvantageous application of the community's resources.

The measures are seen to be temporary. But when it comes to protective duties this insight is slow to come, may never come. The vested interests stubbornly oppose any return to the previous and more advantageous channels of industry. All the fallacies and catch words are brought into play: domestic labor against foreign labor, foreign invasions of our markets, the money lost in payment for imports, and so on. Sometimes it happens that by a lucky stroke an industry is called into being which proves able to hold its own even without the prop of a duty—the case of protection to young industries, of which more is said in the next chapter—and then this exceptional outcome is declared to be the common one. All experience proves that when protective duties are once imposed it is extremely difficult to get rid of them, even tho they were originally advocated as mere emergency measures. At the start, they may be more "wasteful" than other ways of making employment or less so; but either way they are more likely to be retained indefinitely.

§ 4. In the United States by far the most common and most effective argument in favor of protection is that it makes wages high or enables wages to be high. With many persons it is an accepted article of faith that American wages can be kept high, and the American standard of living can be maintained, only if there is protection against the goods made by the cheaper labor of other countries.

With this belief goes another closely similar: that free trade may be advantageous between regions which have the same general range of wages—the same "standard of living"—but is harmful to a country of high wages when carrying on trade with one of low wages. Between different parts of the United States, or between the United States and Canada, or between Great Britain and Germany unfettered exchange, it is said, may be permissible. But not so when the United States and Germany confront each

other in the exchange—least of all, if a country like Japan or China stands on the other side! This fear of universal levelling rests on ignorance or misunderstanding of the causes that lead to the differences between countries in money wages, in prices, in general prosperity. There is here the same ignorance and misunderstanding as in the argument from pauper-labor competition. None put forward in favor of protection are more specious and widely held, none are more fallacious.

Evidently the argument is not of universal application. How could there be any exports at all, if lower wages always gave the foreigner an advantage? As much is exported (virtually as much) as is imported. The exported goods are made by laborers who get high wages in the United States; yet these goods, so far from being undersold in foreign countries, are themselves underselling those of the foreigners. The explanation is simple: the effectiveness of labor in the exporting industries is great, and therefore high wages and low prices coexist. And that effectiveness is the *cause* of the high money wages; and these wages, again, may or may not be accompanied by high prices of the domestic commodities which are outside the realm of international trade. This whole subject cannot be understood except in connection with the principle of comparative advantage. In those industries in which the United States has a comparative advantage in effectiveness, high wages can be paid, and yet low prices accepted, with profit to the employing capitalists. In those in which there is no such advantage, the current high wages cannot be afforded. In this latter class, tho labor be as effective as in competing foreign countries and tho the industries in that sense are well adapted to the country, they encounter the difficulty that other industries are still better adapted, yield still larger returns, and set up a prevalent high rate of wages which these less advantageous industries cannot sustain.

Of course it is true that, when once industries which possess no sufficient advantage have been established under the shelter of protective duties, high wages can be maintained *in those industries* only by the continuance of the duties. This sort of situation—the existence of industries dependent on duties—was historically the occasion of the protectionist argument about wages. Wages have

always been higher in the United States than in other countries. Before a protective system was adopted, it would have been absurd to say that they were due to any such system. When new industries are called into existence by protection, they must, of course, in order to secure their workmen, pay the same wages as are generally prevalent; and once they are established, it can be maintained with reason that high wages to their workmen are dependent on protection. As long as the workmen remain in those industries, the high wages they receive are so dependent.

The free trader argues that if the duties were given up and the protected industries pushed out of the field by foreign competitors, the workmen engaged in them would find no less well-paid employment elsewhere. Presumably they would betake themselves to the exporting industries, in which labor is advantageously applied. The protectionist answers that there would then be "overproduction" in those industries—that more goods would be produced, prices would be lower, and then wages lower. No, replies the free trader—there would be more goods, but not lower prices or lower wages. For there is a new demand for these exportable goods. The new exports must be paid for by imports; there is a new foreign "market," replacing the lost domestic "market." Goods are imported which were formerly made by protected industries. The eventual result, says the free trader, is that more workmen will be turned to the advantageous industries, and more goods will be exported in exchange for more imports; there will be higher wages (in terms of commodities) all around within the country, resulting from the more productive direction of its labor.

In all this reasoning, the free trader is right. There are some further questions concerning the effect of the supposed change on money wages, which will be presently considered;[1] but these do not affect the essentials of the argument. Of course the reasoning applies only to the long-run course of events. It assumes that labor (and capital, too) will shift from a less profitable to a more profitable industry; that when a protected industry is deprived of support, and those engaged in it are confronted with the alterna-

[1] See Chapter 37, § 1.

tive of either accepting lower wages or quitting, they will quit and go to better-paid occupations. Any such process of transition is difficult and trying. When carried out on a very large scale—say by the sudden abandonment of a protective system under whose shelter many industries have grown up—it may cause for the time something like disaster. The extent to which existing industries are in fact dependent on protection is commonly exaggerated by both its advocates and its opponents; but none the less the question of vested interests is a very troublesome one. It may be deemed better, on the whole, to let things stand, or change them very slowly and cautiously, rather than incur the disturbance and damage of a radical change. But all this does not affect the question of principle, which is not squarely presented unless we ask what would have been the best policy from the outset.

The question of wages—to anticipate for a moment—is at bottom one of productivity.[1] The greater the productivity of industry at large, the higher will be the general level of wages. There are very intricate problems as to the precise nature of this connection, and as to the shares of the total which go respectively to wages, interest, business profits, and rent. Under certain contingencies, it is conceivable that protective duties will affect the process of sharing, and so will influence wages otherwise than thru their effect on the total product. But these are rare contingencies and are negligible for the discussion of the main problem. Whatever lessens a country's general productivity tends to lower wages. Protection aims to restrict the geographical division of labor; in doing so, it ordinarily turns industry into less advantageous channels (possible exceptions will be considered in the next chapter). Ordinarily it lowers general productivity, general prosperity, general wages.

§ 5. One phase of the wages argument appears in the proposition, much heard in the United States of late years, that duties should be so adjusted as to "equalize cost of production" between this country and foreign countries. This has been propounded as a "scientific" solution of the tariff problem. When the labor cost of a commodity, it is said, is higher in the United States, let a duty be imposed sufficient to enable the domestic producer to meet his

[1] See below, Chapters 41, 42, Vol. II.

foreign competitor on terms of equality—and then let them fight it out.

It needs little reflection to show that such a policy, consistently followed, means the complete wiping out of all the advantages from international trade, nay, the wiping out of international trade altogether. The greater the disadvantage of a country in producing a given commodity, the more labor must be given to producing it, and the higher will be the expenses of the employers. In proportion as the efficiency or productivity of labor is less, more must be paid out in wages to secure the greater amount of labor required per unit of output; then "labor cost" is so much higher; and duties must be made correspondingly high if the labor cost is to be equalized. Any commodity, however unsuited to the industrial possibilities of a country, can be produced in it if only its price is made high enough; and by keeping out foreign competitors, there is no limit (short of the possible extinction of demand) to the rise in price. If the principle of equalizing cost were consistently carried out, we should exert ourselves most strenuously to promote by high duties the domestic production of an article according as we gain most from its importation and lose most by its domestic production. No doubt, the persons who propose the principle would probably refrain from pushing it to its logical conclusion. They would shrink from clapping on duties high enough to cause lemons to be grown in Maine, or (to use Adam Smith's familiar illustration) grapes in Scotland; tho all this could be done if labor costs were unflinchingly equalized. They think only of the commodities for which the domestic disadvantages are not glaring. But the difference is only one of degree. There is no rational reason for saying that a disadvantage in labor cost—that is, a disadvantage in industrial effectiveness—of 20 per cent should be offset by a protective duty, but that one of 50, 100, 200 per cent should not be so offset.

One thing is to be said in favor of the notion: duties should certainly not *exceed* the rates necessary to "equalize labor cost." If they so exceed, there is the possibility that a domestic monopoly may levy additional burdens on the consumers. This possibility arises if competition among the domestic producers is not free.

As will presently appear, no special benefits to the protected producers accrue, and no monopoly profits are derived, if domestic competition keeps prices down to the level of expenses of production. Where there is a possibility of monopoly and of abnormal profit to the protected capitalists, it is not unreasonable to say that, if they must have protective duties, these should not be greater than suffice to enable the industry to be carried on. But it is absurd to urge that the proposal, even in this form, is a "scientific" solution of the protective question. It simply amounts to saying that protection should not be carried to the point where it may foster monopoly.

§ 6. The strength of the general presumption against protection will be made clearer by a consideration of the working of protective duties in greater detail.

When a duty is imposed on a commodity, its price usually rises by the amount of the duty. It does so usually, but not necessarily; and even in those cases where this normal result is to be looked for it does not always come at once, but often only in the end. Strictly, the result is to be expected only if the commodity is produced under free competition and under the conditions of constant return.[1] Ordinarily a duty, like any tax on a commodity, increases by so much the expense of getting the article to market. The amount of the tax or duty must be added to the price charged

[1] If a commodity is produced under the conditions of diminishing or of increasing return, the case is obviously different. Under diminishing return, a tax per unit of quantity tends to check consumption, lessen production, lower marginal cost, and so increase price by less than the amount of the tax. Conversely, under increasing return, a tax, by lessening consumption, tends to raise marginal cost and so to increase price by more than the amount of tax. A tax on a monopolized article works out its results thru the principles of monopoly value; and it is quite conceivable that such a tax, in the case of an article for which the demand is highly elastic, will cause little rise in price and will be borne chiefly by the monopoly producer. All these possibilities, however, appear in the case of internal taxes quite as much as in that of import duties. They present no special problems in international trade; they are part of the theory of value. Moreover, they are not often of much practical consequence. As intimated in the text, the usual case is, in the long run, that of constant return. The most important qualification of the general reasoning probably is to be made for articles subject to a quasi-monopoly of good will or trademark, where the producers, tho they have no permanent or unqualified monopoly, make unusual profits for a considerable time, and can possibly be deprived of a part of these profits thru the operation of a tax. Compare what is said below, Chapter 72, Vol. II.

the consumer if the producer is to get his usual return. But a rise in price has its effect on demand. Very likely the same quantity cannot be sold at the higher price. The producer, none the less, may not be able to lessen the supply with any promptness; he may have a large plant committed to making the particular thing. For a while, therefore, price may be raised by less than the amount of the tax; conceivably it may not be raised at all. Only as supply is slowly adjusted to the new situation will normal conditions be regained and the price raised so as to recoup the producers and dealers for their increased expenses of production. Hence it is true that a duty on imports, and indeed any tax on a commodity, may fall for a while on the producer, foreign or domestic; while yet in the end it falls with its full weight on the consumer.

So long as the commodity continues to be imported, this rise in price brings a tax, but no national loss. It is true that the consumers are in effect deprived of so much of their incomes; but what they lose, the public treasury gains. Taxes are presumably levied for useful public purposes; they do not stand for waste. If the needed revenue had not been got by customs duties, it would have been got in some other way, and the same amount of tax would have been levied on the public.

Suppose, however, that after the duty has been imposed, domestic producers supplant the foreigners. They charge higher prices than the foreigners did; they *must* charge higher prices, in order to get a profit. If they could bring the commodity to market at the same price as the foreigner, there never would have been any importation. The fact that the domestic producers did not enter the field before the duty was imposed, shows that they are under a disadvantage.[1] When they are stimulated by the duty to enter the field and sell their article at a higher price than the imported one had previously cost, the consumer pays the tax in precisely the same way as if the article continued to be imported—that is, in the shape of higher prices. Only, there is in this case no revenue to the public treasury. The extra price stands for so much bonus to the domestic producers, to enable them to maintain themselves in a disadvantageous industry. And it represents

[1] But see what is said in Chapter 37, § 2, on protection to young industries.

so much national loss. In most discussion of protective duties, at least in the United States, the common assumption is that the creation of a domestic industry, supplying a commodity which was previously imported, represents so much gain. Strictly, the reverse is the case. The payment of duties on continued imports brings no loss; the loss arises when the domestic supply supplants the imports, and duties are no longer paid.

Hence where the principle of free trade is consistently followed, a customs duty on an article is accompanied by an internal tax of the same amount on the domestic product. Then the combined taxes operate solely to bring in revenue, and have no effect on the direction of industry within the country. Such was the system which Great Britain long followed with complete consistency. Her customs duties were limited to a few articles of general consumption, such as tea, coffee, cocoa, sugar, beer, spirits, tobacco. On beer and spirits, an internal tax was imposed at the same rate as the customs duty. The other articles were such as would not be produced within the country; the duties on them were of a purely revenue kind. Sometimes, in popular discussion, it is said that the imposition of any duties whatever is inconsistent with the principle of free trade. Obviously, this is a mistake; it is only the imposition of duties that cause a substitution of domestic products for imported that conflicts with the principle.

When a customs duty operates to bring into existence a domestic industry, the domestic producers do not make unusual gains; that is, they do not if the commodity is brought to market under competitive conditions. Very likely those who take the initiative in producing the article make unusual profits on the first imposition of a duty. In time, however, profits will fall to the normal level, and at that normal level prices will be higher than foreign prices only if a real disadvantage handicaps the domestic producers. In other words, nobody gains, and the community loses—the loss consisting in its paying more for the protected article than it would have had to pay without the protection.

Where there are not competitive conditions,—where there is a monopoly, complete or partial, permanent or temporary—the domestic producers may make unusual gains. To the extent that

they do so, another item enters into the account. There may not only be some national loss, but in addition a shift of income from one set of persons to another set. The commodity may be produced at higher expense within the country, and may have to sell on that account for a higher price than if imported. It may sell for a price still higher, because the domestic producers are in a position to keep out competition and make unusual gains. It may even happen that the imposition of a duty enables domestic producers who are under no disadvantage at all, and who could bring the article to market as cheaply as the foreigners, to form a combination and exact a price higher than the competitive one. In such a case there is no national loss at all. What the consumers lose the monopolists gain.

Naturally enough, this last-mentioned case is precisely that in which protection is least popular, tho in a sense least harmful. Where the protected producers make no unusual gains, the system is supposed to work not unfairly. But the direct robbing of Peter to pay Paul, which seems to appear in case of monopoly, strikes the popular imagination at once and leads to indignation; even tho, on critical consideration, it appears that Paul gains merely what Peter loses and that the community as a whole is no worse off. The more distant consequences on general industrial effectiveness which strict economic analysis brings out are within the ken of comparatively few persons.

The ease with which popular feeling can be roused against a monopoly has led to the frequent allegations by opponents of protection that it breeds monopoly. It was once remarked to a congressional committee of investigation that "the tariff is the mother of all trusts," and the aphorism became the text of many free trade sermons. Its truth is limited. The causes of combination are deeply rooted in the industries of modern times. They are found mainly in the development of production on a great scale; a tendency so far-reaching cannot be ascribed to a single external cause.

It is true, however, that protective duties have sometimes brought combination more easily and at an earlier date, and sometimes have increased the gains from it. This is likely to be

the case where the situation is ripe for consolidation within the country, but not ripe for international consolidation—a stage of development not uncommon, especially in the United States during recent years. The tendency to combination, strong and far-reaching tho it is, does not work out its results automatically, irrespective of favoring causes or legislative influences, or international complications. Protective duties in the United States have been at times during the last generation a favoring cause. Tho the trust problem is in its essence very different from that of protection—a graver problem, and of far larger social consequence—the two may interlace.

Just as protective duties may bring unusual gains to some capitalists, if these can keep out competitors, so they may bring exceptionally high wages to some workmen, on the same condition of keeping out competition. This is commonly less easy for the workmen; but it is not impossible, at least for considerable stretches of time. It is feasible most of all in occupations of the handicraft sort, calling for special acquired skill and not subjected to the machine processes. Such was until recent times glass blowing. Certain kinds of glass, especially window glass, called for the services of the blowers, whose trade was not easily learned. They had a tight union, restricted entrance to the trade, and maintained exceptionally high wages. The employers in this industry also combined; so that there was a double monopoly of capitalists and workmen, promoted by very high import duties. The two favored sets alternately quarreled and joined forces, with the advantage in the end, as usual in such cases, to the employers. Here, as elsewhere, new inventions came in, and the application of machinery tended to deprive the handicraft workmen of their special advantage. But so long as the old conditions remained the tariff system may be said really to have kept up wages—not wages of workmen in general, but those of a limited group. And in such cases, as in that of government industries,[1] workmen in general are likely to regard this advantage to a small group with approval, even tho it may mean higher charges to consumers and to the great body of the workmen as consumers. Anything that means high wages to any

[1] Compare what is said in Chapter 66, § 5, Vol. II.

set of manual laborers finds favor with the labor leaders and doubtless with the dumb rank and file also; partly from mere clannish sympathy, but mainly from inability to distinguish between the causes that bring advantage to all and those that bring advantage to a favored few only.

CHAPTER 37

PROTECTION AND FREE TRADE (*Continued*). SOME ARGUMENTS FOR PROTECTION

§ 1. Protective duties, by their effects on general money incomes, may bring more advantageous terms of international exchange.—§ 2. Protection to young industries. Applicable in the main to manufactures only. Difficulty of gauging its success in specific cases.—§ 3. Political considerations, illustrated by the case of shipping subsidies.—§ 4. Subsidies; much resorted to in the 17th and 18th centuries: less so in the 19th; more so again in the 20th.—§ 5. Social considerations may tell against manufactures, but not necessarily so. The controversy in Germany; Agrarstaat *vs.* Industriestaat. The argument as to the failure of food supplies.—§ 6. Peculiar dependence of England on international trade and on exports. Possibility of strengthening her position as exporter by agreements with colonists and by threats of retaliation.—§ 7. Growth of protection during the 20th century. Still greater growth after the war of 1914–18. Autarkie in Germany; in the British Commonwealth.—§ 8. Economic effects of protection in the United States; impossible to measure accurately, but certainly exaggerated in popular discussion.—§ 9. Conditions under which manufactures would maintain themselves without protection. Effect of machinery in connection with comparative costs. Concluding remarks on the working of protection in the United States.

§ 1. THE simpler aspects of the protective controversy have been considered in the preceding chapter—those which bring out most strongly the case for free trade. They tend to show that the increase in price due to a protective duty represents a net loss. But there are ways in which the loss may be offset. The consideration of the various possible modes of offset brings out those arguments for protection which have some validity.

First there is a possible influence on the terms of international exchange.[1] The first effect of a duty is almost always to lessen imports. Even if it be a purely revenue duty, it will lessen them; the rise in price will cause a decline in consumption, unless demand happens to be quite inelastic. If the duty is protective and

[1] In the sense in which that phrase was used and explained in Chapter 35.

operates to stimulate domestic production, the decline in imports will be more certain and greater. Hence the movement of specie will be into the country. Then will ensue the train of consequences (always supposing the flow of specie to be considerable and continued) already familiar to the reader. Prices and incomes rise within the country, and fall in foreign countries. Exports in time begin to be checked, as the prices of exported articles rise; imports are stimulated, as the prices of imported articles fall. The length of this period of transition, and the extent of the change before it comes to an end, depend on the play of reciprocal demand. If the commodities exported from a country are of a sort insistently demanded in foreign countries; and if, on the other hand, the commodities which it imports are not such as to be consumed more largely as their prices fall—then the change may be considerable. Eventually equilibrium is reëstablished; exports diminish and imports increase until payments again balance. When this stage is finally reached, the country that imposed the duty will have higher money incomes and higher prices. The higher incomes will be of no benefit so far as domestic purchases go, since within the country prices have risen in the same proportion. But they will be of advantage in the purchase of things imported.

In such a case, there is a balance of loss against gain. The consumers lose as purchasers of the protected articles, that is, of those made at home under the influence of the duties; but they gain as purchasers of things that continue to be imported. Even if the particular articles subjected to the duties are completely shut out, there will remain imports of other articles. Thus in the United States protective duties have served to prohibit completely the importation of many manufactures; but tea, coffee, sugar, tropical articles of all sorts, sundry raw materials, some finer manufactures, have continued to come in. All these, if the reasoning of the preceding paragraph holds good, are got in reality more cheaply because of the duties. It is true that some of the things imported, being still subject to duty, are absolutely raised in price; but for this advance there is a full recompense in the revenue received by the public treasury and in the relief (presumably) from other taxes. But even these imports are not raised in price by the full

amount of the duties—there is some offset because foreign prices in general have fallen, and domestic money incomes have risen.

How far is reasoning of this sort applicable to the concrete facts? Precisely to the same extent as the general reasoning on the distribution of the gains from international trade. How difficult it is to verify this in detail has already been shown. Take the case of the United States during the half-century following the Civil War, when a system of high protective duties was steadily maintained. Thruout the period a whole series of other factors influenced international trade, some in one direction some in the other. The protective system, in so far as it restricted imports, made for gain in the terms of exchange, the high tariff contributing something toward a higher range of money incomes. How far the gain from this source served to offset the loss from the fact that the domestic commodities were produced at higher cost and sold at higher prices is impossible of calculation. In any event no such possibility is reckoned with at all in the popular controversy. Most people who try to persuade the public on one or another side of the tariff question reason only about what is "good for business," about employing labor, higher prices to consumers, extortionate monopolies. Even the simpler questions really involved, as to the general effects of the geographical division of labor, they perceive but vaguely; the more intricate ones here considered are quite beyond the understanding not only of the average man but of the average writer on protection.

It is obvious that all countries could not play this game. No one of them has a monopoly of imposing import duties. A condition of mutual grasping and recrimination may be imagined, in which each country tries to get from the other all it can, with the eventual result that, while some advantage accrues to one among them in the form of high money incomes, considerable loss to that country and to the rest is entailed from the curtailment of the advantageous division of labor. Commercial strife has come perilously near this state in modern times; but the immediate object held in view by the combatants has never been that of getting some of the imports cheaper. The motives and objects have invariably been of a semi-mercantilist sort: to check imports and yet

to market more and more exports. Reciprocity movements are a compromise resulting from this familiar sort of contest.

§ 2. The argument for protection to young industries points to another way in which the main argument in favor of free trade can be fairly met and the initial loss from protection offset. The gist of it is that an industry really advantageous for a country may be prevented from arising because of ignorance, lack of experience, all the obstacles that impede success in unfamiliar undertakings. Stated in another way, the argument is that while the price of the protected article is temporarily raised by the duty, eventually it is lowered. Competition sets in, it is said, and brings a lower price in the end. The free trader asks, why any need of a duty if the domestic producer is really able to sell at a lower price than the foreigner? The protectionist answer is that the reduction in domestic price comes only with the lapse of time. At the outset the domestic producer has difficulties and cannot meet foreign competition. In the end he learns how to produce to best advantage, and then can bring the article to market as cheaply as the foreigner, even more cheaply. Most persons who use this second form of the argument (alleging the eventual reduction of domestic prices) are but dimly aware of its identity with that for protection to young industries. But the two arguments are one and the same, resting on the premises of temporary obstacles and eventual success.

The theoretical validity of this argument has been admitted by almost all economists. The question is how far and under what circumstances there is ground for applying protection with prospect of this good result. The argument was first used (in such a way as really to make an impression) in the United States during the earlier part of the nineteenth century, when this country was in the transition from dominantly agricultural and commercial conditions to the stage of modern manufacturing. It was carried from the United States to Germany by its best-known advocate, Friedrich List, who applied it to Germany in her transition during the middle of that century from semi-medieval to modern conditions. The United States was then a "young" country, and Germany, tho an old country, had manufacturing industries that were

young so far as modern ways were concerned. In both countries there was force in the contention that manufactures with machinery, power, large-scale operation, were certain to arise in any case, or at least had an advantageous opportunity; and that the process of transition and growth could be made easier, and a beneficial result could be reached at an earlier date, by a temporary handicap on the developed competitors of older countries. England of course was the country then in the van, against which such shelter was sought.

List and the other more moderate advocates of nurturing protection said that duties for this purpose should be moderate and should be temporary. They should be moderate—not to exceed say 25 per cent—because, if the domestic industry was at a great disadvantage in the beginning, there was little prospect that it would ever reach independence. They should be temporary—not to endure more than twenty or thirty years—because in the end, by supposition, the domestic industry would not need them, and ought to be able and willing to face foreign competition. It was further added that agricultural commodities and raw materials give no field for this sort of protection. Their geographical distribution is determined chiefly by unalterable physical conditions. Only in manufacturing industries can the legislator have a prospect of encouraging young industries with good results.

These limitations on the argument are reasonable; more particularly the exclusion of agricultural articles. The government can do much to promote efficiency in agriculture; but chiefly by diffusing education, improving the conditions of tenure, promoting science. There are respectable arguments, as will presently appear, for duties on such articles; but they are of a very different kind from this one, which looks to promoting eventual cheapness. The United States long levied protective duties on wool, but never with any prospect of getting wool cheaper thereby. Germany and France levied duties on grain, as England did until 1846 and after 1931; but in none of these cases was there any prospect of thereby securing domestic supplies more cheaply.

The other limitations seem also reasonable; but in actual experience it is not so clear that they must be observed in order to

secure the desired result. Not only moderate duties, but very heavy ones, may set things going, and eventually lead to an independent domestic industry. Of this possibility the recent history of the silk manufacture in the United States supplies an illustration. A duty of 60 per cent on silks was imposed during the Civil War (1864). The object at first was revenue. Then a domestic industry grew up; and the duty was maintained, even increased (especially in 1897). Competition became active, and great improvements were introduced. The silk manufacture has indeed been the last of the textile industries to be adjusted to the machine processes; but this development seems to have been promoted in the United States by the establishment of the industry under the shelter of protection. It is certain that advances in manufacturing methods have taken place; it is probable that some branches of the industry, tho not all, have reached the stage where the fabrics can be put on the market as cheaply as they can be imported. Nor is it inconsistent with this outcome that the domestic producers still clamor for protection. They are simply in the habit of doing so. Most business men know very little outside the immediate range of their business. If foreign competition has been long shut off by a high duty, they are ignorant of its possible effects; and if there is a proposal to permit it again, they object on general principles even tho they are quite able to hold their own. The protective system, especially when exaggerated stress is laid on it thru party politics, begets an abject fear of all foreign competition. Notwithstanding this common attitude of the domestic producers, it is quite possible that the object of protection to young industries was in fact attained; tho the only certain way to ascertain this would be to remove the duties and let the domestic producers meet the foreigners on even terms.

While it is possible that protection to young industries may be successfully applied where advantages in production rest not on natural grounds but on acquired skill, it is extremely difficult to say how far there is a probability of such success. The question is part of one much wider—the general causes of the advance of the arts. Economic history shows that the spread of the various trades and manufactures in different countries has taken place by no

"natural" process, and that "artificial" factors, such as governmental encouragement, the emigration of skilled artisans, the social and political organization of a country have been of large, often dominant, effect. It would be absurd to apply to the conditions of medieval and early modern times a theory of natural advantages and of settled differences in comparative advantages. On the other hand, the lesson of history seems to be that other modes of encouragement have been more effective than protective duties; such as rational education, free industry, abatement of social barriers, promotion of invention by patents and trade-marks. In very modern times—with the wide diffusion of industrial education, the ease of communication, the technical press, the eager search for all ways of investing capital at a profit—the argument for protection to young industries would seem to have lost much of its force. None the less, possibilities still exist, as in the case of the silk manufacture just cited. Unfortunately the decisive test—eventual removal of duties—is one which domestic producers are likely always to oppose; and so long as their opposition is successful it will be difficult to ascertain in any particular case whether the community ultimately gets a real gain sufficient to offset the initial loss.

§ 3. Political considerations are often urged in favor of protective duties.

The most conspicuous illustration is afforded by shipping. In the days of wooden vessels, a merchantman was not so very different from a man-of-war, and at all events training in handling the two was the same. Moreover, a merchant marine was an effective auxiliary in times of war. The first of these reasons is less important in our day, when steel battleships have intricate and highly specialized machinery of their own. The second is perhaps as important as in former days. A modern navy needs an elaborate complement of scout ships, supply ships, colliers, not to mention transports. A large mercantile marine supplies these, or at least aids in meeting the suddenly increased need of them which arises in time of war. If, to use Adam Smith's phrase, defense, for that matter, aggression?, is more important than opulence, it will be worth while to promote a mercantile marine, even tho it cannot do

its work so cheaply as foreign shipping. It might even be economical to subsidize a merchant marine, under conditions which assure the availability of the merchant ships in time of war; this course being very possibly cheaper than that of hurriedly creating an auxiliary fleet when war breaks out.

The protection of shipping, viewed simply as a matter of the adjustment of a country's productive forces, presents no new question of principle. If foreign ships can carry goods more cheaply than domestic ships, let them do it, says the free trader. There is no wonder-working magic in having your own ships. They exist simply to carry goods; and the same grounds which hold for letting the foreigner produce and sell goods to you, if he can do it more cheaply, hold for letting him transport goods for you, if he can do this more cheaply.

The only economic peculiarity in the shipping situation is that the usual method of protection, by duties, is not here available; at least not for shipping engaged in foreign trade. True, a system of preferential taxes can be elaborated, tonnage duties being made higher on foreign ships than on domestic, or on goods imported in foreign bottoms. But this sort of discrimination invites easy retaliation. The domestic ships so favored must in due time go to foreign ports, and in those ports they in turn may meet the same sort of hostile treatment. Not only may they do so, but they certainly will. Retaliation of this sort has been universally applied. Hence all countries have found themselves compelled to enter on reciprocity arrangements for vessels engaged in the direct trade between them, and have agreed to treat domestic and foreign vessels on the same terms. For shipping in the foreign trade, the only available protective policy is that of direct subsidy. Coastwise shipping—from one port to another in the same country (including colonies)—is of course not subject to this limitation, and here protection can be applied without hindrance. Most countries which maintain protection in any form apply it to the coasting trade by excluding foreigners once for all.

National pride and prejudice, which have been important factors in promoting the growth of protective feeling, have been particularly so in regard to shipping. The Stars and Stripes have

disappeared from the seas: here is the most effective popular argument in favor of shipping subsidies. Vaguely associated with this are the arguments in favor of a merchant marine as a means of supplementing a fighting navy. This combination of sentiment, military glory, and serious political considerations belongs outside the strict domain of economics. As a matter of material welfare the drift of all rational economic thinking is against subsidies to shipping. Soberly considered, merchant ships are but implements for promoting the division of labor, and the Stars and Stripes waving over them are cause for patriotic exultation only if the ships are made and handled to the real advantage of their own country and of other countries as well.

§ 4. While subsidies—direct payments to domestic producers —are the one form of protection available to the shipping trade, in other industries there is a choice between subsidies and import duties. During the seventeenth and eighteenth centuries, the period of unblushing mercantilism, subsidies were freely applied over a wide range. In the course of the nineteenth century, they were gradually dropped, few remaining at its close. Their disappearance was the result in part of the administrative difficulties inevitable in guarding them from abuse, even more of the general current against protective measures. It is one of the most curious features of the great changes in all these matters during the post-war period of the twentieth century that direct subsidies or bounties again came to be paid on a large scale in the most varied parts of the world,—not only on the Continent and in Great Britain, but in countries as different as Australia and British India. In Great Britain, once the tenacious upholder of free trade principles, bounties were given on articles of the kind for which the free trade reasoning had been supposed to have the greatest validity, such as agricultural products (wheat and sugar beets). This change, like so many others among the unexpected disruptions brought on by the war of 1914–18, rested on interests, prejudices, trains of reasoning in which the economic aspects were either glossed over or deliberately set aside; an aspect of the situation of which more will be said presently.

Bounties or subsidies are a less popular and hence a more vul-

nerable method of protection than import duties. Import duties, while in essentials they come to much the same thing as bounties, can be defended by a host of persuasive tho fallacious arguments. The direct payment of money to a favored industry presents in unmistakable form the question whether it is really worth while thus to tax the community. From the free traders' point of view, this very simplicity is an argument in favor of using in all cases bounties and subsidies rather than import duties. The general recognition of the need of taking account of the repugnance to bounties is shown by the fact that they are not often granted without limit of time; they are usually for a stated period only, after which they cease either at one stroke or by gradual steps. It is rare that any provision for automatic cessation or reduction is made when protective duties are imposed.

It is this difference between subsidies and protective duties—as to probability of permanence—which signifies most. As regards lasting effects, the economic difference is in the ways in which the lessening of the national dividend is worked out. The real enduring consequences are the same. In the case of duties, the ultimate outcome is that the prices of the taxed articles are higher than they would otherwise be; while yet the government gains no revenue. Bounties, on the other hand, cause no rise in prices of the subsidized articles, even tho the industries are such as could not be carried on without them. The bounties are a drain on the public revenue and merely serve to offset the inability of the assisted industries to produce the articles as cheaply as they could be had thru importation. This is the gist of the free-trade reasoning—the economic reasoning. In the end, no individuals gain, but the community loses. All else is but qualification and explanation.

The first effect in both cases, however, is that somebody gets gains at the expense of somebody else; and in both it is the stage of initial gains which is most conspicuous and bulks largest in popular discussion. In both cases, again, the stage of high gains may last for a considerable period. The more imperfect competition is, the longer will the gains last; if there is a well-intrenched monopoly, indefinitely long. Whatever the conditions—pure competition, imperfect competition, monopoly—what we hear loudest

from the free traders is talk about ill-gotten gains, plundering one set of people at the expense of another, robber barons; while the protectionists avow that they are increasing the demand for labor, helping the workingman, fostering domestic industries, keeping the money in the country. In all of which there is—as regards the short-period consequences—a modicum of truth on both sides, yet little understanding of the heart of the question: whether government by either sort of legislation can succeed in directing industry into more productive channels and so in raising the total national income of goods and services.

§ 5. Considerations as to general social soundness are supposed by some to strengthen the case for free trade, by others that for protection. But it is doubtful whether a strong case can be made out on such grounds either way. It is said by the protectionists that diversified industry brings social and educational advantages and that a community whose occupations have a very narrow range will be deficient in intelligence and adaptability. In view of the degree of industrial diversity which is certain to appear under any circumstances in a modern country of advanced civilization, this sort of vague allegation has no probative force. Free traders have argued on the other hand that a diversity of industries secured by the promotion of manufactures at the expense of agriculture is more likely to bring bad social and political consequences rather than good ones. Manufactures mean large-scale production, concentration in comparatively few hands of management and of ownership, dependence of workmen on wages by hire, increasing inequality. They mean, too, crowding in cities, and the temptation to exploit women and children. In the earlier part of the nineteenth century arguments of this sort were much used in the United States against protection. They were not without weight; they may not have quite lost their weight. The soundest parts of our American nation are in those regions of the North where agriculture is still the dominant industry. But after all the mode in which an industry is conducted and the character of the people engaged in it are more important than the nature of the industry itself. The workingmen of the English manufacturing districts in Lancashire, Yorkshire, and Scotland are better social stuff than

the agricultural laborers of eastern Germany and probably even than most of the peasant proprietors of France. Protection and free trade are minor factors as compared with the diffusion of education, the general range of intelligence, the distribution of wealth and income, the demarcations of social classes, political and industrial freedom.

A special application of social and political arguments was made in Germany in the opening years of the twentieth century, combined with reasoning of a strictly economic sort. There the controversy was between the advocates of the *Agrarstaat* and of the *Industriestaat,*[1] the former being in favor of duties on grain and other agricultural products, the latter opposed to them. To the former—the protectionists—dependence on foreign countries for indispensable foodstuffs always seems to entail evils and dangers. An agricultural population, or at least a population with a due proportion settled on the land, is thought to be better social material than one mainly engaged in manufactures. A great development of manufactures, moreover, and a dependence on foreign markets for disposing of the products, bring uncertainty. Hostile tariffs, or the loss of the advantage in production on which the exportation rests, may put an end to the trade and endanger the established industries. Finally—and here the crux of the arguments was reached—the relations between the European manufacturing countries and the oversea countries which developed during the last quarter of the nineteenth century were regarded as essentially temporary—temporary, that is, compared with a nation's life history. The supply of food, and especially of wheat, from the United States, Argentina, Canada, rest on methods of cultivation which could not be permanently maintained.[2] The continuous use of the soil for the same crops can be kept up only so long as new land remains available. Sooner or later—and it will be soon, say these protectionists—the virgin lands will all be occupied; and then a conserving cultivation, with varied crops, must come. Meanwhile, population in the new countries increases

[1] The German word "Industrie" means "manufactures." The English "Industry" commonly refers to all the operations of production. But "industrial," whether used as adjective or noun, refers to manufactures.

[2] See what is said in Chapter 42, § 5, Vol. II.

rapidly, their own consumption of foodstuffs becomes greater, their economic situation becomes steadily less favorable to the exportation of grain and the like.

Hence—so the argument ran—those old countries in which great manufactures develop, based on an exchange of the manufactured products with imported food, must face the possibility, nay the probability, of an eventual revulsion. Food will no longer be obtainable by importation. The manufacturing population must then go back, in part at least, to the land. But this population, under the stimulus of plentiful employment and cheap food, will have become large, and an endeavor to support it at home will meet all the obstacles of diminishing returns from land. The example of England is held up as a warning. Her great population, which the country's own resources cannot possibly supply with food and materials, is necessarily dependent on foreign trade, and must be constantly uneasy lest trade with other countries may fail.

There is much validity in this train of reasoning. As put forth by careful thinkers, it admits the *prima facie* loss from protection. In the present generation food is indeed got cheaper by foreign trade, and the exchange of manufactures for food is for the time being advantageous. Some of the ardent protectionists hesitate in this sort of admission, as people commonly hesitate and minimize in such concessions to their opponents; but the admission must be made. It must be admitted also that the process of checking the growth of manufactures by making foodstuffs dear is a trying one. It is a sacrifice to the future which in the present generation may be unpopular. But where the sentiment of nationality is strong and the welfare of coming generations is prized, such sacrifice may be willingly made.

To go into all the details of the controversy on *Agrarstaat* and *Industriestaat* would pass the limits of this book. The free traders aver that in a country of great extent and diversified climate like Germany, no such extreme development of manufactures as in England is to be looked for; that the probability of failure of supplies from food-exporting countries is exaggerated; that if there comes eventually a check to the exchange of manufactures for food, it will be by no sudden disastrous halt, but by a gradual

process to which industry and population can adjust themselves; and finally that, for the present generation, the burden of import duties is heavy and that the chief beneficiaries are a small knot of large landed proprietors. The main economic argument of the protectionists, as to the future failure of food supplies, raises a question difficult in many directions—namely, how far it is wise to go in the restriction of immediate satisfactions for the sake of a distant and more or less uncertain future. Shall we now for example husband our coal supplies, which we know to be limited? Or shall we use them freely according to present needs, partly indifferent to the distant future, partly trusting to possible discoveries and improvements for other sources of heat and power? Shall the Germans (and the English, too) maintain a policy of free trade and of dependence on distant countries for food and materials needed now, without speculating too anxiously upon the continuance of these supplies in the uncertain future? It is easy to err in endeavoring to provide too solicitously for coming generations. Such are some of the large problems which the protective controversy presents in a country like Germany—problems which give fair ground for differences of opinion, and involve considerations much weightier than those usually put forward by protectionists in the United States.

§ 6. A similar phase of the tariff controversy developed in England. True, the steps towards a manufacturing nation (*Industriestaat*) had been irrevocably taken, and the question was as to the best means of remaining with safety and prosperity in this far-developed stage. It would seem at first sight that here a policy of free trade alone is tenable. Yet the reaction against it appeared in England also, and was not without the support of effective arguments. These arguments, so far as they are of weight, turn mainly on the expediency of reciprocity arrangements.

In the preceding pages it has been said more than once that exaggerated importance is commonly attached to a country's exports. For a country in England's situation, however, there is substantial ground for watching the exports with special care and perhaps with some anxiety. They are the means for obtaining indispensable imports. The alternative of producing the imports at

home—of turning the labor and capital from making the things exported to making those now imported—hardly exists. England must import; and in order to import, she must export. Hence everything which lessens the market for exports must cause concern. Among the ominous possibilities is the imposition of protective duties elsewhere. It is a matter of large consequence for England to maintain in other countries an open market for herself. Hence the advocacy of imperial federation or imperial preference duties as a means of inducing the colonies to relax, if not to give up, their duties on English goods; and hence the advocacy of duties on foreign goods in England, as a means of chaffering with other countries in negotiations for the reciprocal reduction of tariff barriers. In England, as in all countries, the vulgar fallacious arguments in favor of protection play a large part in the popular controversy: such as increased employment for home labor, support of domestic industry, tribute to foreigners in payments for imports, and so on. But these arguments are more insidiously dangerous in England than anywhere else. That country depends for its very existence on manufacturing industries which are able to face the competition of the world. If once these industries, one and all, cannot face foreign competition—if they really must lean on protection against foreigners—her future is dark. The only solid ground for advocating duties on all the manufactured products is that some industries are still in the van and that an all-inclusive system would enable the diplomats to higgle more effectively for the admission of their products into other countries. And the only ground for preferential arrangements with the colonies is to induce them to admit English goods with no duties or with duties lower than those imposed on non-British goods.

In general, a retaliatory policy, that is, the levy of duties on imports in retaliation for duties elsewhere on a country's exports, makes the economic situation not better, but worse. If Germany levies duties on English goods, the advantages from the division of labor between the two countries are lessened by so much. If England then levies duties on German goods, those advantages are lessened by so much more. If, indeed, one takes the Mercantilist view of foreign trade and assumes that its chief object is to procure

a market for the exports, then retaliation and reciprocity assume a different aspect. Then a country becomes always intent on increasing its exports and always uneasy at increasing its imports; and then it will perhaps consent to admit some imports more freely only if tempted by a bait of selling some exports more freely. So long as this state of mind exists, there is at least a possibility of securing an eventual relaxation of restrictions by first imposing restrictions.

What may be the substantial grounds for expecting a real extension of international trade thru reciprocity treaties, it is difficult to say. Adam Smith remarked that this matter was not for the economist but for "that crafty and insidious animal called the statesman or politician." The staunch free traders aver that other countries will go their way undisturbed by retaliatory duties or preferential offers, or will make concessions that are only nominal; and that the initiating country herself will suffer at once from her own restrictions, and in no way gain in the end. On the other hand, it must be admitted that the Mercantilist notions persist with extraordinary tenacity. The immense majority of persons think of a reduction of duties not as a gain to their own country but as a favor shown to the foreigner; and conversely they think of tariff reductions by foreigners as the opportunity to sell more goods abroad and profit thereby.

§ 7. The growth of protection during the closing years of the nineteenth century and the opening years of the twentieth was a remarkable phenomenon, in view of the weight of rational opinion against most of the arguments commonly advanced for it. During the generation following the repeal of the English corn laws in 1846, the indications seemed to be that free trade, or at least a great relaxation of customs barriers, would extend over the civilized world. But in the decade 1870–80 the current began to turn the other way. Country after country moved toward protection, and even in England, the home of free trade, signs of reaction appeared. The protectionist movement is explicable on various grounds. The growth of nationalist feeling was one important cause. Protection seems, to most people, a "national" policy, and in fact is so, in the sense of causing exchanges to be made within a

country rather than between countries. The principle of free trade has a certain cosmopolitan flavor, and assumes as well as promotes a spirit of peace and good will among the nations. Another cause was the admitted need of a thoro reconstruction of economic theory. This promoted skepticism as to free trade, which was one of the cardinal doctrines of the older school; altho no part of the system of the older economists has stood the test of time and criticism better than their reasoning about international trade. Still another cause was the competition of oversea countries with the agricultural producers of the Continent. The landed interest there, formerly indifferent or hostile to duties, joined in the demand for protection against underselling foreigners. At all events, during the generation preceding the war of 1914–18 a wave of protection succeeded the previous one of free trade.

After the war of 1914–18, the wave of protection mounted higher and higher, and free trade was quite swept away. In part this arose from the same causes as those of the pre-war period. Economic theory gave more and more attention to imperfect competition, and to the play of demand rather than to any steadying influence of supply. It cannot be said that the essential conclusions concerning international trade, as stated in the preceding pages, were seriously impugned; but their exposition could not be so confident and dogmatic. As regards the whole wide range of state action for directing the channels of industry, a less confident tone about international trade was the result of the growing belief that there must be more of conscious and deliberate control of economic activity; more planning all around, not merely in the international distribution and diversification of industry but also in monetary systems, the ordering and steadying of domestic industry. Much of this was brought about by the extraordinarily severe depression which set in after 1929, the distracting confusion of the industrial and monetary situation, the resort to any and every means which might somehow promise to promote stability. In the field of international trade the most important element was the intensification of nationalist feeling. That led to an even more stringent protectionist policy than before the war in the United States and in continental countries, and to the definitive adoption by Great

Britain of a bargaining and "quid pro quo" policy. The extreme form of protectionism appeared in Germany; the advocacy of autarchy (Autarkie), which came from people of the most various sorts,—mystics, idealists, fanatics, political and economic leaders. It called for isolation in every country and complete economic independence, above all for Germany. Such a consummation obviously was not feasible unless a country had a large and diversified area. The proclaimed ideal was both effect and cause of the demand for political changes. Its realization might be deemed not impossible for the United States, or for a newly created state or confederation in the middle of Europe. The British Commonwealth might conform to it after a fashion. Even tho it could not give promise either of complete free trade within the range of that huge and unwieldy Commonwealth, or of complete isolation from non-members, it might bring about such moderation of protection in the several constituents of the Commonwealth as would maintain a great part of the advantages from free trade while still placating the demand for friendly trading with your countrymen and aloofness from others.

In all this there was more froth and foam than in the pre-war period, more loose talk and superficial thinking; and also more attention, even tho not of a wiser sort, to other than economic considerations. So far as the economic aspects go, there was little to say beyond what had been threshed out again and again. It was a time when aspects other than economic were forced on men's minds more and more; both the relations of social classes to each other and those of nations and peoples. These matters of political, social and ethical philosophy are in part quite within the range of economics, but in good part go beyond it. So far as international trade is concerned, the reasonings and conclusions of our science are based on the assumption of continued peaceful dealings. Their trend is against war, against the blocking and disruption of trade. It is toward amity, good will, coöperation in all the dealings of men with each other.

§ 8. In the United States a severely protective tariff was maintained for half a century after the Civil War. The financial exigencies of the war caused high duties to be levied, and in sub-

sequent years these were retained. A rigid and all-inclusive system of protection grew up, and persisted without serious modification (barring a brief reaction in 1894–97) during the post-war period, —indeed with a marked accentuation in 1922 and in 1930.

The economic effects of this system it is impossible to follow empirically. We have seen that its effects on the terms of international exchange are so interwoven with those of other factors that no unraveling is possible. Even more baffling is the task of following or measuring its effects on general prosperity. The protectionists, on this subject as on the rate of wages, have preached and protested that all good things come from their tariff. Such talk results naturally from the exigencies of partisan conflict and the need of simple arguments for the mass of voters. So loud and persistent has been the talk that for many persons, even for many who are not unintelligent or uneducated, it has become an article of faith that the prosperity of this country rests on the protective tariff. Yet there is no greater delusion. A multitude of factors explain our general welfare—vast resources, a far-spread division of labor within the country, a free, active, and intelligent population. Has not this North American region been for several centuries, under all sorts of economic and political conditions, the envy of the world? But to trace in detail the part played by any one factor in promoting or retarding the enviable outcome is well-nigh impossible. Certain it is that, so far as the tariff is concerned, we must rely chiefly on general reasoning. The first and obvious effect of protection is to turn industry into less advantageous channels; and there is, in my judgment, no good case to rebut this general conclusion and to establish a balance of gain from such a tariff system as the United States has had since the Civil War.

Yet it should be said that on many articles the duties were but nominal. These are the articles which were made as cheaply within the country, and (competition being active) were sold as cheaply. The mere imposition of a duty does not raise prices. It does so only if a foreign supply is cut off and a more expensive domestic supply is thereby induced, or a domestic monopoly fostered. The extent to which manufacturing industry in the United States is dependent on the tariff system is vastly exaggerated by the pro-

tectionists. One would suppose, from their doleful predictions, that not a chimney would smoke but for the tariff. In fact, the United States is certain to be a great manufacturing country under any conditions. So much is assured by its resources of coal and minerals and by the ingenuity and enterprise of its people. Its comparative advantage is by no means confined to agriculture. But this same consideration indicates that the free traders went too far in ascribing ill effects to all the parts of the protective system. It did not change the course of industry so greatly as their charges implied. The country would be as prosperous, would have industries as diversified, without a high tariff as with it. The quantitative effect of a protective tariff is commonly exaggerated in all countries, not in the United States only.

§ 9. The conditions on which depends the maintenance of manufactures in a country like the United States deserve a moment's consideration. The usual cause of advantage in manufactures is better machinery and methods. Take the case of the shoe manufacture, which has been cited as one of our efficient and independent industries. Shoes are not imported; they are beginning to be exported in considerable quantities. The Americans have taken the lead in the invention and perfection of machinery for making them. But machinery can be bought or copied. The Germans, perhaps, can copy it, and then, working it with cheaper labor, can undersell the Americans. This is often true of the Germans, or at least was so; they have been good imitators, tho slow originators. It is said that American steel skates, devised and perfected in the United States, were copied to the smallest detail in Germany, and then, being made there with cheaper labor, were imported into this country. This sort of imitation is not usually possible; since for the operation of machinery a force of intelligent and skillful mechanics is often as necessary as the machinery itself, and is much more difficult to secure. But the thing is possible, if not always at least in many cases; and the more so if machinery becomes automatic. The salvation of the industry then is, in a country like the United States, incessantly to improve machinery. Constant progress is the condition of maintaining the comparative advantage. Once the same methods—that is, the same effectiveness

of labor—prevail the world over, then the country where wages are lower can sell cheaper.[1]

It is commonly said that the United States is likely to have an advantage in those manufactures where machinery is much used. This is true; but the real explanation is not often given. The mere use of labor-saving machinery does not give an advantage. Machinery represents only one way of applying labor. It is the use of labor-saving machinery to a greater degree or in a more ingenious way that enables the output to be comparatively cheap even tho the wages of laborers are high. In those industries which are adapted to the machine processes, American labor is *likely* to be more efficient. Which those industries are, cannot be settled by any rule. The march of invention is irregular. Sometimes Americans take the lead, sometimes Englishmen, sometimes Germans or Frenchmen. It is proverbial that Americans have a more than creditable record in this sort of competition; and the economic corollary is that they do well to confine their manufacturing activity to those industries in which they seem able to keep in the van.

In some cases in the recent history of manufacturing industry in the United States, it is to be admitted that this process of getting the lead seems to have been promoted by protection. That is, protection to young industries has been successfully applied. The object has been attained by a rude, blundering, expensive method; but we must grant that attained it has sometimes been. The silk manufacture has already been cited as an example. Possibly the iron and steel manufacture presents another. But this latter case is more doubtful, because the question always arises whether such an industry, already established on a moderate scale in the country, would not probably have grown to independence under any circumstances. The steady increase and thickening of population and the growing scarcity of free land tended in any event to bring about the development of other than agricultural industry. The great streams of immigration and the altered conditions of labor

[1] This holds true, that is, of any one industry. If *all* industries had the same methods and the same efficiency the world over, there would presumably be no differences in wages. International trade would then cease. Cp. Chapter 34, § 4.

supply thereby brought about strengthened still more this tendency. The tariff system, even where it may seem to have acted in the way of protection to young industries, often merely quickened development which would have come soon enough without it.

In this review of the tariff problem nothing has been said of some of its more obvious bad aspects—the pressure of interested producers to obtain measures favorable to themselves, the contributions of a semi-corrupt character to party chests, the log rolling by which each legislator strives in the general scramble to secure duties that will be of benefit, or at least will be thought of benefit, to his own constituents. The tendency in popular government for each representative to press the real or supposed interests of his special constituents is the greatest evil of democracy. It has been experienced to the full in tariff legislation. But it appears in many directions, in things good as well as in things doubtful—in education, harbor improvements, the postal service, public control of railways and other industries. Something of the sort must be faced whenever the State undertakes to direct and regulate matters of immediate economic consequence. We should give attention chiefly to the general outcome, under such working conditions as the existing state of political machinery makes possible; and from this point of view the question of protection also must be judged.

REFERENCES ON BOOK IV

On the foreign exchanges, an older book, dealing with the main principles, is G. J. Goschen, *Theory of the Foreign Exchanges* (1861; revised later editions.) More recent is G. Claire and N. Crump, *The A B C of the Foreign Exchanges* (10th edition, 1936), chiefly on the details of the mechanism in England. A. C. Whitaker, *Foreign Exchange* (2d edition, 1933), is full and detailed on the mechanism, especially in the United States, and also deals with the larger questions of principle. The effects of dislocated exchanges are well considered in C. R. Whittlesey, *International Monetary Issue* (1937). The same subject is treated with a wealth of information in S. E. Harris, *Exchange Depreciation* (1936).

On international trade, outstanding among the expositions of what is called the "classic" theory are the chapters in Book III of J. S. Mill's *Principles of Political Economy,* and Book III of Marshall's *Money Credit and Commerce* (1923). Still later is Taussig, *International Trade*

(1924). An able statement, much compressed and not always easy to follow, is in Bastable, *Theory of International Trade* (4th edition, 1903). A systematic treatment is J. O. Yntema, *A Mathematical Reformulation of the Theory of International Trade.*

More recent books, giving a larger share of attention to such subjects as international investments and dislocated exchanges, are G. Haberler, *Theory of International Trade* (1936); P. J. Ellsworth, *International Economics* (1938). B. Ohlin, *Interregional and International Trade* (1933), covers, as its title indicates, a wider range than earlier books. J. Viner, *Studies in the Theory of International Trade* (1937), not a systematic treatise, is noteworthy for its discussion both of the development of theory from the 17th century to the 20th and of the theoretic problems of the later date.

Notwithstanding the mass of literature on Free Trade and Protection, no book covers this one subject satisfactorily. The best treatment is—as one would expect—in the books on international trade at large, such as those of Haberler and Ellsworth, just mentioned. An excellent discussion of the tariff issues at large and also of the English conditions of 1900–31 is in *Tariffs: The Case Examined* (1931) by a committee under the chairmanship of Sir W. Beveridge. On the economic effects of protective duties in the United States I venture to refer to two books of my own, *Some Aspects of the Tariff Question* (3d edition, 1931) and *The Tariff History of the United States* (8th edition, 1931).

लाल बहादुर शास्त्री राष्ट्रीय प्रशासन अकादमी, पुस्तकालय
L.B.S. National Academy of Administration, Library
मसूरी
MUSSOORIE 105563
यह पुस्तक निम्नांकित तारीख तक वापिस करनी है।
This book is to be returned on the date last stamped

दिनांक Date	उधारकर्त्ता की संख्या Borrower's No.	दिनांक Date	उधारकर्त्ता की संख्या Borrower's No.

330.1
Tau
4th ed.

105563
अवाप्ति सं.
ACC No. ~~ID-179~~

वर्ग सं. पुस्तक सं.
Class No..... Book No.....
लेखक
Author...Taussig, F. W.....
शीर्षक
Title...Principles of economics.
.....

निर्गम दिनांक | उधारकर्ता की सं. | हस्ताक्षर

330.1
Tau
4th ed.

~~I.L.~~

LIBRARY
LAL BAHADUR SHASTRI
National Academy of Administration
MUSSOORIE

Accession No. 105563

1. Books are issued for 15 days only but may have to be recalled earlier if urgently required.
2. An over-due charge of 25 Paise per day per volume will be charged.
3. Books may be renewed on request, at the discretion of the Librarian.
4. Periodicals, Rare and Reference books may not be issued and may be consulted only in the Library.
5. Books lost, defaced or injured in any way shall have to be replaced or its double price shall be paid by th