ECONOMIC THOUGHT
OF THE
TWENTIETH CENTURY

ECONOMIC THOUGHT OF THE TWENTIETH CENTURY

CLAUDIO NAPOLEONI
EDITED AND TRANSLATED BY
ALESSANDRO CIGNO

Economic Thought of the Twentieth Century

CLAUDIO NAPOLEONI

edited, expanded, translated and introduced by
ALESSANDRO CIGNO

A HALSTED PRESS BOOK

1972
[1963]

JOHN WILEY & SONS
New York

The original Italian edition was published in 1961 by Edizioni RAI Radiotelevisione Italiana and the second edition in 1963 by Giulio Einaudi editore under the title 'Il Pensiero Economico del 900'.

This edition was first published in 1972 by Martin Robertson & Company Ltd., 17 Quick Street, London N1 8HL.

Published in the U.S.A. 1973 by the Halsted Press, a Division of John Wiley & Sons, Inc., New York.

ISBN: 0 470 63010 8
Library of Congress Catalog Card No.: 72–12304

Printed in Great Britain

EDITOR'S PREFACE

This brief history of modern economics originates from a series of talks given by Claudio Napoleoni on the Italian radio between 1960 and 1961, and first published in 1961 under the title of *Il pensiero economico del 900*. Because of its origins and length the present book cannot go into detail and avoids technical jargon as much as possible, but an attempt is made to put over the gist of the theories rigorously enough to avoid misleading the reader. The line of thought underlying the selection of topics and the link with theoretical developments before 1900 are made explicit in the Introduction. The theories left out are mainly those which are not concerned with the functioning of the economic system as a whole, but with particular aspects of it however important: the theories of consumer and investment behaviour are the most notable victims, and monetary theory comes in only as part of wider theories.

As a textbook, this book could serve as an introduction to the study of the original sources in a course on the history of economic thought, or it could be used alongside the usual textbooks of a theory course to help the student see the link between the various topics and place them in historical perspective. It should also be of value to the non-specialist who wants to become acquainted with the kind of problems with which economists exercise their minds; in Italy it is popular not only among students and teachers of economics, but also with the many who have an interest in the social sciences in general.

The English edition is more than a translation. The major change with respect to the original is the addition of an Introduction and a new chapter on growth theory which I have contributed. Chapter 8, which in the Italian edition covers a number of topics having in common only their mathematical nature, here deals exclusively with modern equilibrium theory and related areas; the rest of the material originally in it has been partly omitted and partly moved to Chapter 4. The original footnotes have also been replaced with a bibliographical note at the end of each chapter, which contains

the basic references and a guide to further reading. Some other minor adjustments had to be made mainly because the last Italian edition is now nine years old.

The translation is based on the second Italian edition and, without being literal, attempts to remain as close as possible to the original; at only a few points does the English depart from the Italian text for the sake of clarity. I am much indebted to my wife Katy and to Mrs Patricia Newnham for their invaluable help in rendering Italian prose into something close to modern English.

A. C.

CONTENTS

INTRODUCTION

The best way of introducing this book is, I believe, to talk about David Ricardo. To start a history of twentieth-century economics with an early nineteenth-century author may seem odd, but the fact is that the problems posed and left unsolved by Ricardo have haunted economists ever since, and this book ends with the attempt by Piero Sraffa to get around Ricardo's difficulties. A history introduced by Ricardo and concluded by Sraffa may not be everyone's idea of modern economics, but at least it makes a coherent story.

The focus of Ricardian analysis is on the agricultural sector, where production takes place by applying labour to land and can only be increased by extending cultivation to less fertile soil. As workers need to be paid at frequent intervals, while output comes at the end of the cropping season, production is only possible because of the existence of 'capital', a wage fund accumulated by capitalists in return for a profit—their remuneration for 'waiting'. In the short run the wage is equal to the existing wage fund divided by the given labour force, profit is the difference between the product of the marginal worker and his wage, and the average rent on the land is the difference between the product of the average worker and that of the marginal one. In the longer run the labour force varies with the size of the wage fund, which in turn depends on the rate of profit. But labour productivity diminishes as the margin of cultivation is extended, therefore a situation will eventually arise where the wage has fallen to bare subsistence level and the rate of profit is at the minimum acceptable by capitalists. At that point capitalists have no further incentive to accumulate and the economy grinds to a halt. These conclusions apply even if there are other industries besides agriculture, because wage and profit rate must be the same in all parts of a competitive economy; so the agricultural rate of profit determines the distribution of income for the economy as a whole. The division of income between consumption and saving is also determined because all subsistence wages are obviously consumed and the rest, profits and rents, is assumed to

be entirely saved.

The problem with Ricardo's approach—shared by Smith and Marx—is that the rate of profit must be calculated *before* prices can be determined. This would be possible if agriculture produced only one commodity—say wheat—which was also the only form of consumption and therefore used to pay wages. Thus profits and capital would both consist of wheat, and the agricultural rate of profit could be calculated in physical terms as the ratio of the two. Competition would then determine the prices of all products in such a way that the rate of profit would be the same in all industries. If there were more than one wage-good, the method would fail because the rate of profit could not be calculated in physical terms even in agriculture, and we would need prices in the first place in order to calculate profits and capital in value terms. This means, in effect, that the theory can only stand if the economy produces only one final good—used both for consumption and for accumulation as capital—and all the others are intermediate goods which disappear in the course of production.

Another objection to Ricardo's approach is that consumer preferences do not come into the picture at all, so that consumption can only be determined by the minimum subsistence requirements of the population. Thus the whole *raison d'être* of the productive system is not to satisfy the needs of the community to the highest possible degree, but to reproduce itself as fast as it can by subtracting from consumption as much as possible of the final product: a foolish aim—inevitably frustrated by the fixity of land—which will take the economy straight towards a stationary state. This and other theories of economic stagnation are examined in Chapter 7.

For all these reasons the 'classical' approach was abandoned towards the end of the last century and replaced with what is generally known as the 'neo-classical' approach. The neo-classical view is that capital is not a wage fund but a collection of durable goods required by further production: given a stock of these goods, natural resources, the state of technical knowledge and the preferences of the consumers at any point in time, competition (the 'invisible hand') simultaneously determines the quantities and prices of all products, and the prices for the services of natural resources and capital goods. The rate of profit or return on capital is then simply the ratio between the price for the service of a capital-good and the price of the capital-good itself. The Ricardian difficulties therefore seem to have disappeared and the consumer becomes the centre of the universe: within the constraints imposed by natural resources and technology, consumer preferences ulti-

mately decide all the results of economic activity. Furthermore, rent, wage and profit are now the competitive prices for the services of land, labour and capital—the just reward for everyone's contribution to the common welfare—and the economic system is no longer the stolid and cruel machine described by the classics, but the reflection of a just society.

Neo-classical logic has permeated much of modern economics and borne abundant fruit. This way of reasoning allowed Walras to specify the concept of price as an index of scarcity (see Chapter 1), Pareto to find the meaning of optimality in economics (Chapter 2), Robbins to define economics as the science of rational choice (Chapter 2), and—rather ironically, as neo-classical theory was intended as a rationalization of capitalism—it even laid the foundations for the rational management of a socialist economy (Chapter 10). Neo-classical optimism was rather shaken in the 1930s, when it was claimed that in reality competition is far from perfect and Chamberlin and Joan Robinson showed that the performance of the economic system is inefficient. Worse still, it was shown that when competition is only among a few the neo-classical method cannot give definite conclusions (see Chapter 4). Keynes, on the other hand, questioned the postulate that the economy is ruled by consumers' decisions and, by recognizing the role of speculation and businessmen's expectations, demonstrated the possibility of prolonged under-utilization of resources, particularly labour (see Chapters 5 and 6). Neo-classical theory—which Keynes confusingly insisted on calling 'classical'—thus appeared to be better suited to showing how the economy ought to work than how it actually works in a capitalist society.

But the neo-classical approach also hides a logical difficulty similar to that of Ricardo (see Chapter 1). If capital is regarded as a collection of durable products as in Walras, there is in fact no way of ensuring that the rate of return is the same on each individual capital-good; on the other hand, if capital is seen as the cumulative *value* of resources committed to further production as in Wicksell, prices would have to be known in advance of production in order to calculate the value of capital—which is not possible. This difficulty prompted Pareto to ignore accumulation and growth altogether for the sake of arriving at a logically consistent theory of static equilibrium. Schumpeter, on the other hand, argued that profit has no place in a static economy as profits are the result of change, which gives the innovators a margin over the current cost of production; the rate of profit and, consequently, the market rate of interest on loanable funds can therefore be determined only

in a growing economy, where it depends on businessmen's readiness to try out new methods and explore new markets (see Chapter 3).

Interest in the problem of accumulation revived after the Second World War, when growth became the overriding preoccupation of governments and academic economists. There are two main strands in this line of research. One is concerned with the underdeveloped economies, the choice of the technique of production that will allow them to reach an acceptable level of consumption in the shortest possible time and the problems involved in getting the growth process off the ground and generating enough saving to finance the accelerated accumulation of capital required for rapid growth (see Chapter 11). The other is concerned with the characteristics of the growth process in developed capitalistic economies, and the ways in which private and public decisions can alter them (see Chapter 12). It is mainly in connection with this second strand of research that the need for a logically consistent theory of capital and income distribution was felt, and the classical and neo-classical difficulties re-emerged.

A classical solution to the problem of determining simultaneously a unique rate of profit and the prices of all goods had already been found by Von Neumann (see Chapter 8). He showed that when an economy is expanding at the highest speed permitted by technology the rate of profit is equal to the rate of growth, and prices are determined by the proportions in which goods have to be produced in order to allow the economy to grow at maximum speed. To obtain this result Von Neumann had to ignore consumer preferences and non-produced resources, and treat consumption as the input required by the production of labour. But many goods growing in a constant proportion are the same thing as one (composite) commodity, and a consumption determined by production requirements can only be at subsistence level: we are therefore back to Ricardo—only without land.

The neo-classical approach to growth was also to assume that the economy produces only one composite commodity, but consumption was seen as the end of production, not vice versa. Accordingly, the labour force was assumed to grow at its own pace, independently of production, and consumers were allowed to choose between the immediate consumption of the composite commodity or its accumulation for further production and future consumption. One of the implications is that the growth rate of the economy is determined not by technology but—as Harrod pointed out—by the increase in the labour force and in labour efficiency; another implication—derived by Solow and Swan—is that the choice of the

technique of production, the rate of profit and the distribution of income are determined by consumer preferences. This approach was criticized on Keynesian grounds by Kaldor, because it does not recognize the role of businessmen in determining the rate of accumulation and technical progress and relies on the assumption of perfectly competitive markets to transmit the consumers' decisions to the producers. Another criticism was that, outside a Von Neumann path, it is not legitimate to regard the composition of the composite commodity as fixed; therefore, the rate of profit cannot be calculated in physical terms and we need prices before it can be calculated in value terms—Ricardo's problem. The neo-classical theory of growth and what can be called its neo-Keynesian counterpart are reviewed in Chapter 9.

A measuring stick which does not vary with prices was eventually invented by Sraffa (see Chapter 13). He devised a special kind of composite commodity—the 'standard commodity'—having the peculiar property that the individual goods composing it are the same and are combined in the same proportions as those required to produce it. The rate of profit can therefore be calculated in physical terms for the standard commodity and will rule throughout the economy, but the proportions in which individual goods are actually produced do not need to be the same as those in which they enter the standard commodity: Sraffa shows that prices are determined in such a way that the rate of profit calculated in value terms is the same as the one calculated in units of standard commodity. Ricardo's contradiction and the fundamental problem of capital and distribution theory are therefore solved. Sraffa's contribution, which appeared in 1960, was intended as the prelude to a critique of current economic theory and, indeed, it has provided ample ammunition for attacks on the prevalent neo-classical approach, but it has not yet given rise to a fully worked out alternative.

Alessandro Cigno
January 1972

Chapter 1

THE SITUATION AT THE BEGINNING OF THE CENTURY

1. *The theory of equilibrium*

At the beginning of this century economic science appeared to be dominated by the theory of equilibrium, which was formulated between 1870 and 1900. Most subsequent developments have been either further elaborations of equilibrium theory or criticisms of it, which is why a proper understanding of the history of twentieth-century economics must start with a full explanation of this theory. An explanation is even more necessary because implicit in the theory of equilibrium is a concept of economic activity and, consequently, of economic science which, after being defined about 1930, has become the keystone of subsequent theoretical elaboration.

Equilibrium theory developed from the work of the major economists at the end of the last century: Menger in Austria; Jevons, Edgeworth and Marshall in Britain; Walras in France; Pareto and Barone in Italy; Clark and Fisher in the United States; Wicksell in Sweden. As modern critics have recognized, the differences between these economists are differences of detail, and there is a unity of standpoint and theoretical development in their work which justifies considering them as co-authors of a single doctrine.

For the purposes of this book it will be useful to examine in particular the form—one of the most complete and rigorous—which the theory of equilibrium took with Walras between 1873 and 1877.

2 *The Walrasian theory of general economic equilibrium*

Walras' theory is called the theory of *general* economic equilibrium because it is concerned not only with the equilibrium of single economic units (households and businesses) considered separately, but above all with the position of equilibrium reached by the whole economic system. Walras' problem is therefore the following: given

certain initial quantities of productive resources, a certain technology and the preferences of the individuals, what then determines the quantities of goods produced and exchanged, as well as the prices at which the goods are exchanged, in the general equilibrium configuration—i.e., in that situation where the positions of equilibrium towards which the individual units tend are achieved? It is important to note that the type of economy that Walras has in mind in constructing his theory is, at least in intention, absolutely general: that is, not conditioned by elements specifically relevant to this or that social system. In other words, Walras is concerned with characteristics common to every economy, without reference to the particular institutional framework in which it belongs; but he only partially succeeds in this, because the essential elements of his construction, and the various categories and functions which appear in it, are drawn directly from the capitalist reality.

Walras' reasoning can be fully understood if one bears in mind the concept of social wealth which he takes as the starting-point of his theoretical treatise. This concept of wealth must be examined carefully, for it not only influences Walrasian theory, but is also at the bottom of nearly all subsequent scientific speculation. By *social wealth*, Walras means 'all things, material or immaterial . . . that are *scarce*, that is to say, on the one hand, *useful* to us and, on the other hand, only available to us in *limited quantities*'. There are, therefore, two conditions necessary for something to be defined as wealth: it must satisfy a need and it must be available in limited quantity with regard to this need. Neither useless things, nor those which, while useful, are available in such quantities (like air) as to completely satisfy the need for them, are part of social wealth. The twofold quality of being useful and limited confers three properties on wealth: it is *appropriable*, it is *exchangeable* and '*it can be produced and multiplied by industry*'.

'Pure' economic theory, according to Walras, inasmuch as it is the theory of the determination of prices and quantities produced and exchanged, can therefore be defined also as the theory of social wealth. Walras bases the theoretical determination of the equilibrium position on an accurate classification of the elements which make up social wealth: this classification is important because it enables the various functions and the various types of actions which take place in the economic system to be distinguished. The first major distinction made by Walras consists in dividing social wealth into two categories: capital and income. *Capital* consists of goods which are used more than once, or durable goods. *Income* consists of goods which can be used only once. Capital includes in its turn

three categories of goods: land, personal capital (i.e., the work capacity of the individuals present in the economy) and capital proper (buildings, machines, etc.). Income includes non-durable consumer goods, intermediate goods (i.e., non-durable goods used in production, such as raw materials) and, finally, the services of capital goods, that is, the uses of the three kinds of capital already mentioned.

On the basis of this classification, equilibrium theory can be divided into four stages. The first, called 'theory of exchange', is concerned with determining the quantities exchanged and the prices of consumer goods. The second, called 'theory of production', is concerned with the quantities exchanged and the prices both of the productive services of capital and of intermediate goods. The third, called 'theory of capital formation', is concerned with the quantities produced of capital goods in the narrow sense and with their prices. The fourth stage takes into account the fact that there may be a lag between the time when the final products are sold and the time when intermediate goods have to be purchased. This makes it necessary for the producers to obtain advances on their sales and gives rise to the phenomenon of circulating capital. Each of these phases naturally includes the preceding and represents therefore a further approximation to general equilibrium.

A situation of general equilibrium can be described as follows. Imagine that three categories of individuals are present in the economic system, distinguished by the nature of the 'capital' at their disposal: *landowners* owning land, *workers* who possess personal capital, and *capitalists* who own capital proper. Each individual, as an owner of capital, is able to offer on the market the productive services of the capital itself. Besides these three categories of individuals there is yet another, that of the *entrepreneurs*—those who buy productive factors (whether services of capital goods or intermediate goods) on the market and, on the basis of a *given* technology, combine them in productive processes which produce consumer goods, intermediate goods and capital goods proper. The buyers of these products, in the case of intermediate goods, are the entrepreneurs themselves. In the case of consumer goods and capital goods they are the same persons who supplied the productive services in the first place and who now, spending the income gained from the sale of these services, and on the basis of a *given* structure of preferences, buy, as consumers, consumer goods and, as savers-investors, the capital goods made available by the productive process. Given the quantities of capital goods of each kind initially available to each individual, it is then a matter of estab-

lishing what are the quantities produced and exchanged of each type of good, and what are the prices of these goods. This is done by supposing that conditions of *perfect competition* exist on all the markets which make up the economic system; that is, the demand for, or supply of, each particular good by each individual should be small enough, in comparison with the total amount exchanged on the market, so as not to influence the market price. This means, in other words, that while prices are determined by the combined actions and reactions of all the individuals, each individual, taken by himself, must accept the market prices as *data* that cannot be modified.

The logical process by which an equilibrium position is reached can be described as follows. Suppose that, purely by chance, a certain price system exists. Such a system consists of consumer goods prices, intermediate goods prices, the prices of productive services, the prices of capital goods and the rate of interest which relates the price of the service of each capital good to the capital good itself. In the presence of such prices, each individual tries to act in such a way as to obtain a position of maximum satisfaction or 'utility'. The individual owners of capital goods will try to sell the services of these capital-goods in order to obtain an income that will compensate them for the sacrifice that this entails. They will then try to distribute this income in the most satisfying way between consumption and saving. They will also try to spend their consumable income on the various goods in such a way as to maximize their utility and will use their savings to acquire capital goods in such a way as to maximize the income that these capital goods will bring in the future. The entrepreneurs will buy the factors of production, whether productive services or intermediate goods, and will combine them so as to produce at minimum cost. They will decide which goods to produce and in what quantities in order to maximize their profits. This combination of reactions to the given set of prices thus determines a set of quantities supplied and quantities demanded that will have to be matched on the market. Of course, since the price system has been supposed to exist by chance, it cannot be presumed that the supply of, and demand for, each good will be equal. Therefore, prices will rise in all cases where demand exceeds supply and fall in all cases where supply exceeds demand. Other demands and other supplies will correspond to each new price system. The process will continue until equality is reached between quantities offered and quantities supplied. Equilibrium will then result from a twofold set of conditions: a *subjective condition*, which consists in each individual pursuing maximum satis-

faction, and an *objective condition* which, through the equilibrium between demand and supply for each market, guarantees that the optimum position reached by each individual is compatible with that reached by all the others.

It is important to point out, from a methodological point of view, that this grandiose construction of Walras, precisely because of the large number of its constituent elements and because of the complex system of relationships between them, cannot be properly expressed by ordinary language: it needs mathematical language. The Walrasian system is, in fact, translated into an imposing system of algebraic equations expressing all the subjective and objective conditions discussed above. Of course, this system was not thought of by Walras as a system of equations that could in fact be solved, but was written in order to bring out the logic of the system and in order to confirm by mathematical reasoning the coherence and generality of his arguments. Others have subsequently examined Walras' problem again and completed it in several points. This question will be dealt with afterwards; at this point only two subsequent developments need to be mentioned. Firstly, it has been shown that the equality between the number of equations and the number of unknowns, which Walras considered sufficient for the accuracy of his treatise, is not sufficient for guaranteeing the existence of economically meaningful solutions. Other conditions are needed for this purpose and this has enabled equilibrium theory to be further investigated in a very interesting way. Secondly, by appropriate manipulating and simplifying, Walras' system has been transformed so as to enable it to be used for actually calculating numerical solutions and this has been of great importance for economic planning. We shall look at both these questions in full later. But one of the most important aspects of Walras' theory must be made clear from the start, and that is its importance for economic planning—an importance that was not at first suspected. To do this, we must first explain the concept of price resulting from the system of general economic equilibrium.

3. *The concept of price in equilibrium theory*

Price, according to current use, is simply the rate at which two goods are exchanged. Generally, one of the many goods present on the market is used as a term of comparison for all the others, and in such a case it is called the 'numeraire'. If the numeraire is, say

gold, then the prices of all goods are expressed in gold, and the rate of exchange between two goods is the same as the ratio between their prices in gold. This elementary idea of price is of course right, but it only touches the surface of the phenomenon. The theory of general economic equilibrium enables deeper aspects of the price phenomenon to be brought out—aspects which are indispensable to the understanding of its true nature.

Firstly, the Walrasian system enables the concept of 'marginal rate of transformation' for the economic system as a whole to be specified. That is, we can determine the additional quantity of a certain good that can be produced when productive resources are liberated by giving up production of a unit of any other good. In other words, a kind of *technological equivalence* is established among the various goods, and through this, by shifting resources from one productive activity to another, it is possible to transform one good into another according to a well-defined law. Secondly, in Walras' system it is possible to define the so-called 'marginal rate of substitution' between any two goods; that is, it is possible to specify what additional quantity of a certain good would be necessary to make up for the loss of a unit of any other good, in order that the satisfaction of the consumer of these goods remains unaltered. It is possible, in other words, to establish a kind of *psychological equivalence* between the various goods from the consumer's point of view.

We also have that, in the equilibrium configuration determined by competition, given any two goods, their marginal rate of transformation in production is equal to their marginal rate of substitution in consumption, and the common value of these two ratios coincides with their relative price as established by the competitive market. The equilibrium prices are therefore the measure of the rates of technical and psychological equivalence between the various goods present in the system. The importance of this idea lies in the fact that, in contrast with the commonsense concept of price, it does not necessarily need the concept of exchange. Which is to say that, given the initial quantities of productive resources, given a certain technology, given the tastes and preferences of the consumers, the price of one good in terms of another can be determined, at least in theory, as the rate of technological and psychological equivalence of the two goods, whether or not an actual act of exchange takes place in a market.

The importance of this idea is obviously enormous; it sheds light on the fact that the theory of general economic equilibrium, born and developed essentially as a theory of the competitive market,

leads to a notion of price which would enable a planned economy to be founded on rational bases, when the problems of computation that go with it are resolved. Walras was not fully conscious of this aspect of his theory, which begins to be clear in the works of Pareto, a major follower of Walras, and becomes completely explicit in the works of another Italian economist, Enrico Barone. Expounding the general equilibrium theory in an article in 1908, Barone gave his work the significant title of *Il ministro della produzione nello Stato collettivista* (The minister of production in the collectivist state). This article was the first explanation of the planner's view of equilibrium theory, and started a tradition later enriched by very important contributions after the Second World War. These will be examined in Chapter 10.

4. *Money and the theory of general equilibrium*

An important aspect of the general equilibrium approach is that it is valid for determining only relative prices and not absolute prices. This is explained as follows. We have already said that the *objective conditions* of equilibrium consist in the equality, for each good, between the quantity supplied and the quantity demanded. Now these equilibrium conditions are not all independent: there is one (any one) which can be logically derived from all the others and, as such, is not a true condition because it does not impose a further constraint on the unknowns of the system. Consider any one of the individuals present in an economic system: according to the Walrasian aproach, he is at one and the same time supplying certain productive services and demanding certain goods, and he will behave in such a way that the total value of goods demanded is equal to the total value of goods supplied. If this is true for each individual, it will also be true for the whole economy, so the total value of products and services supplied will be equal to the total value of products and services demanded (*Walras' Law*). It is then clear that, if there are enough constraints to ensure that demand equals supply for each individual good but one, demand and supply will have to be equal for that one, too—by virtue of Walras' Law.

One of the demand-equals-supply conditions is therefore redundant and must be eliminated, but this results in the number of constraints on the unknowns of the problem being one less than the number of unknowns. Then one of the unknowns has to be regarded

as an 'independent' variable and its value must be fixed arbitrarily. This being the case, let us take any price as the independent variable: all the other prices and all quantities of goods will then be determined by the constraints of the general equilibrium system in relation to that particular price. Now let us ask how all the variables of the system depend on the price taken as the independent variable. A hypothesis characteristic of equilibrium theory will help us here, namely that the quantity demanded and the quantity supplied of each good will not change if all prices vary in the same proportion. This hypothesis can be justified by the two following considerations. Firstly, if all prices change in the same proportion, the 'economic position' of each individual does not change, because the change in price of what he has to buy is accompanied by an identical change in price of what he is selling. Secondly, as the ratios between prices are, by hypothesis, unchanged, the terms of reference for the choices of the various individuals remain the same. It can be demonstrated—and it is anyway self-evident—that, on the basis of this hypothesis (called the 'homogeneity assumption'), in the solution of the equilibrium system each price is directly proportional to the price chosen as the independent variable, while each quantity is constant. This means that the quantities of goods produced and exchanged are independent of the arbitrary values assigned to the independent variable. The rate of interest is equally independent of it, because it is itself already a ratio between the prices of capital-services and the prices of capital-goods. As for the prices, the general equilibrium approach determines them as a function of one of them—and the good to which this independent price refers is called 'numeraire'—and determines them in such a way that when the price of the numeraire varies, the ratios between prices do not change. This approach, therefore, determines *relative prices*, not absolute prices.

In this sense, it can be said that the theory of general equilibrium does not embody a theory of money, since one of the main purposes of monetary theory is to determine the absolute level of prices. This idea, that relative prices, the quantities produced and the rate of interest on the one hand, and the absolute level of prices on the other, are determined by completely separate and reciprocally independent mechanisms, is one which occurs very frequently in modern economic thought before Keynes and is stated with particular clarity by Cassel. Relative prices, the quantities produced and the rate of interest are determined, as we have seen, by the general equilibrium system; absolute prices are determined by a further condition, which can be illustrated as follows. Let us con-

sider a good (gold, for example) which is used as money, and let us ask ourselves what quantity of money is needed by the sum of the individuals in an economy to carry out all the transactions which take place over a certain period, say a year. This quantity of money has a definite relation with two elements: the value of annual production and the velocity of circulation of the money itself (i.e. the number of transactions for which a monetary unit is used, on average, in a year). The exact quantity of money required is equal to the value of annual production divided by the velocity of circulation. With regard to the value of annual production, it should be noted that, from the solution of the general equilibrium system, we can derive both the quantities produced and the relative prices. This means that the *value* of production is determined by arbitrarily fixing one variable. Given this, suppose that the monetary authorities fix the supply of money. Since the quantity of money demanded must be equal to the quantity supplied by the monetary authority, this imposes—in addition to the conditions already imposed by general equilibrium—a further condition which determines the only variable left unknown by the general equilibrium approach, i.e. the 'independent variable'. Once this variable is determined, the absolute price level is obtained, and therefore also the value of production, which will be proportional to the quantity of money.

This theory—called the 'quantity theory of money'—implies a net separation between the *real* aspect and the *monetary* aspect of economic life: the real side determines the relative prices, the quantities of goods and the rate of interest, while the monetary side determines the absolute prices. It is important to point out that the separation between the two aspects is entirely based on the hypothesis that the demand for money arises only from the necessity of carrying out transactions, so that, once the velocity of circulation is established, the quantity of money serves only to determine the absolute level of prices. When, in Keynes (long preceded in this by Walras and by Wicksell in particular), it is affirmed that the quantity of money is one of the elements essential for determining the rate of interest (since money is used for speculative purposes as well as for transactions), then the aforesaid separation will disappear. In the latter approach, in fact, if money is not considered, then the rate of interest is indeterminate and therefore all the variables related to the rate of interest (firstly and immediately the production of capital goods) cannot be determined either. This discussion will be taken up again, more appropriately, with reference to Keynes.

5. *Theories of partial equilibrium: A. Marshall and I. Fisher*

Of the writers who, besides Walras, contributed to the formulation of equilibrium theory, Pareto's contribution at the beginning of this century will be examined later. Here it suffices to mention the work of two economists, Marshall and Fisher, whose theories, in contrast to that of Walras, relate to partial euqilibria; that is, they are concerned with particular markets and not the whole economic system.

The examination of the works of Marshall facilitates our understanding of the new theories of the market which came out between the two world wars. His *Principles of Economics* was for many years the main economics textbook in English-speaking countries, replacing John Stuart Mill's *Principles*, which had dominated the English scene since 1848. Marshall's theory is usually thought of as the theory of partial equilibrium, for he was especially interested in the study of the individual productive unit in a given market, and of the single industry, defined as all the firms producing the same commodity. Naturally, because of this characteristic, Marshall's theory is not so grandiose or complete as that of Walras, but it provides a much richer and less mechanical view of the structure and behaviour of the competitive firm.

Three of Marshall's contributions particularly concern us here. Firstly, the theory of demand which he developed from the results already obtained by Cournot in 1838. Here Marshall set out rigorously the functional relationship between the price of a good and the demand for it, thus clearing away more than a few obscurities which had previously arisen from the confusion between changes within the same demand schedule and changes of the entire demand schedule as a consequence of changes in income or in consumer preferences. Marshall's analysis lays the foundation upon which all the numerous theoretical and statistical studies of demand will develop. Secondly, he lays out the necessary and sufficient conditions for defining a regime of perfect competition. Thirdly, there is his very important analysis of cost and, on the basis of this, his study of the way in which the firm and the industry adapt to market conditions in the short and in the long run.

This distinction between short and long run, although outlined from classical times, is Marshall's contribution. He defined the 'short run' as the length of time during which it can be supposed, firstly, that a firm will not alter the size of its plant and therefore can only change the volume of production within the limits fixed by its available capacity; secondly, that the number of firms composing an industry is given. By 'long run' he meant a length of time during

which both the size of plants and the number of firms are supposed to be variable. As a result of this distinction, it was possible for Marshall to define two types of equilibrium, depending on whether a short- or a long-run perspective is chosen. A firm is in short-run equilibrium if it maximizes its profits with a given plant and the rate of profit can be different from that achieved by other firms and other industries. In the long run the firm is in equilibrium when it maximizes its profits by choosing the most profitable plant size, while firms entering or leaving the industry have brought this maximum profit to the same level in all industries. Walras' general equilibrium is, in the light of Marshall's definition, a long-run equilibrium.

The other 'partial' analysis of great importance, that of Irving Fisher, refers to that particular market where the rate of interest is determined. The problem, according to Fisher, is as follows. When we say that the equilibrium rate of interest is, say, 4 per cent a year, we mean that 104 money units available in a year's time can be exchanged on the market for 100 money units available today. It is, then, a matter of seeing which are the mechanisms through which the market determines this particular price, the rate of interest, which permits the comparison of values at different points in time. Fisher supposes that all the other prices have been fixed: in this lies the 'partial' character of his analysis. The rate of interest is determined on the basis of two principles. Firstly, there is a psychological factor, that is, the undervaluing of the future in respect of the present by the average individual, as a result of an attitude which Fisher calls 'impatient'; it is no other than the circumstance that a good available today is generally preferred to a good available tomorrow. Each individual has his own 'rate of time preference', which will depend on various factors, but mainly on his current income and the way in which he expects it to vary in the future. On quick reflection it will be clear that the lower the current income and the faster it grows over time, the higher this rate is likely to be.

Suppose, as a first approximation, that current and future incomes are given for each individual, and suppose, according to a procedure already followed by Walras, that the rate of interest has been given at random. Then all those whose rate of time-preference is greater than this rate of interest will tend to borrow and will do so up to the point where their present consumption has increased at the expense of future consumption so much, that the rate of time preference has fallen to the level of the rate of interest. Vice versa, those who have a rate of time preference lower than the rate of interest will tend to lend and will do so until these loans have decreased their present consumption in favour of future consumption so much, that

the rate of time preference has increased up to the level of the rate of interest. On the market, therefore, which is supposed to be functioning in conditions of perfect competition, a demand and supply of loans will form in relation to the given rate of interest: if the demand exceeds the supply, the rate of interest will increase, lowering the demand and increasing the supply, and the process will continue until they are equal. A similar process will take place if, initially, the demand is less than the supply. In the equilibrium position, the demand and supply of loans are equal, and the rate of interest is equal to the rates of time preference of all the individuals.

In a second approach to the problem, Fisher introduces another principle important for determining the rate of interest. The hypothesis that the income of each individual is given is abandoned, in consideration of the fact that technology offers many income opportunities and that the use that the individual makes of the resources at his disposal determines the stream of his future incomes. The principle regulating this choice is that the individual will devote his own resources to the use corresponding to the income stream having the maximum present value, calculated by discounting future incomes at the market rate of interest. Once this choice is made, the time path of consumption can be modified by borrowing or lending according to the mechanism described above. The rate of interest that is established in the market in an equilibrium position is thus determined both by the technological possibilities open to each individual regarding the use of resources, and by the rate of time preference of each individual. With regard to the technological aspect of the problem, the most important case by far is the one where the use of resources generating the income stream with the highest present value permits only a small income in the near future and an abundant income in the distant future: such a use is, therefore, an 'investment', which is how Fisher defines the renouncing of an immediate income for a future income. If the lower incomes in the near future are regarded as the 'costs' of the investment, and the larger incomes in the more distant future as its 'returns', then the rate of discount which equates the present value of the returns to the present value of the costs is called the 'rate of return over cost'. An individual will then invest—i.e., move towards alternatives which imply renouncing current income in favour of future incomes —until the rate of return over cost is greater than the rate of interest, that is, until the rate of return on the last dose of income invested has become equal to the rate of interest. There is, therefore, a unique equilibrium value of the rate of interest, which is equal to the marginal rate of return on investment and to all the rates of time prefer-

ence in the position of equilibrium.

Fisher points out that in modern society the bulk of loans are made in order to correct the way in which the income of investors is distributed over time, which, were it not for the possibility of borrowing, would be unbearably distorted as a consequence of investment. Through loans, these distortions are passed on to the lenders and therefore distributed over all the community. Finally, it must be mentioned that technical progress—which Fisher calls 'discoveries and inventions'—plays an essential role in determining the rate of interest: in fact it is from this progress that new occasions for investment arise, and the more intense the progress, the higher the equilibrium rate of interest will tend to be.

6. *Questions left open by the equilibrium theory*

At the end of this introduction to the history of modern economics, it is opportune to mention briefly some problems that equilibrium theory has left open, and which, as we shall see, will be the subject of subsequent thought.

One of the first problems is specifically concerned with Walras' system and is a difficulty within the system itself. Walras' theory is logically coherent in every part except one. This unsatisfactory part of the system concerns capital theory, that is, that aspect of the theory that examines the production of new capital goods as well as the determination of their prices. It does not seem possible, within the framework of the Walrasian system, to demonstrate that solutions exist for the group of equations which refer to the determining of a unique rate of return on the various sorts of capital. In other words, while sufficient mechanisms exist to determine the equilibrium price of the various products through the play of demand and supply, there are not sufficient mechanisms to make the rates of return on the various forms of capital coincide with the unique rate of interest determined by the market.

Suppose, in fact, that the rates of return differ: an adjustment towards equilibrium could take place only through a rise in the supply of services of capital goods with high returns, and a fall in the supply of services of capital goods with low returns. This, in turn, could only come about if the availability of some capital goods increased and that of others fell, which would contradict the hypothesis that the quantity of capital goods is given—a hypothesis fundamental to Walras' scheme because, if it were abandoned, the

system would remain indeterminate. Walras himself was well aware of this when he supposed that new capital goods, produced during the period considered, are unproductive during the same period and only begin to produce in subsequent periods. An indirect but interesting confirmation of this difficulty in Walras' theory is made by Pareto who, after having first explained (in his *Cours d'economie politique*, 1896–7) the theory of general economic equilibrium, including in it a theory of capital identical to Walras', afterwards, following his own line of thought in the *Manuale d'economia politica* in 1906, completely eliminated from the general equilibrium system the phenomenon of the formation of capital, as though he had realized the impossibility of setting up a rigorous theory of this phenomenon within the premises and categories of Walrasian economics.

It is also interesting to note that, in Wicksell's formulation of the general equilibrium theory in 1901, there is a difficulty analagous to that found in Walras. Wicksell's model is different from Walras' with regard to capital. While Walras introduces capital as a collection of durable goods, the quantities of which are assumed to be given, Wicksell, following the Austrian economist Böhm-Bawerk, reduces the capital goods to so-called original factors: labour and natural resources. Such factors would have been invested during the periods preceding the one under consideration, and their current value given as the wages, rents and interests that have accrued to them since they were invested and incorporated in capital-goods until the time when these capital-goods gave rise to production. It is therefore possible to define the total value of the capital invested in the economy as the sum of all these values of the original factors invested in the past. It is this total value which Wicksell assumes as given in each period, in place of the quantities of capital-goods assumed as given by Walras. But, when the *given* quantity of capital is considered as a value and not as a physical magnitude, we meet with the insoluble contradiction that capital depends on those prices (wages, rents and interests) that appear among the unknowns of the problem, so that prices could only be determined by knowing beforehand the value of the available capital, and the value of the capital could only be determined by first knowing the prices.

It must therefore be concluded that the theory of capital (and therefore of interest) has posed great difficulties for the general equilibrium theory, whichever way it is formulated. Nor can such difficulties be avoided by resorting to a theory of the interest rate like that of Fisher which we have examined above, no matter how logically coherent. Such a procedure cannot in fact be adopted

because of the *partial* nature of this theory, which assumes as given all the other magnitudes of the economic system, whereas the difficulty arises within the ambit of a *general* theory purporting to determine *simultaneously* all the economic variables.

The second problem regards the unreality of the hypothesis of perfect competition. This criticism of equilibrium theory constitutes one of the most important components of contemporary economic thought and will therefore be dealt with later.

The third problem concerns the question of growth. Even supposing that the difficulty relating to the formation of capital is not present, it remains that the equilibrium system presupposes that production techniques as well as the tastes and preferences of consumers are given and cannot be modified. In this way, some of the most characteristic phenomena of economic growth are excluded from the system. An economy where production techniques and consumer preferences are unchanging is destined sooner or later to reach a stationary state, where the only possible growth is that of a purely quantitative nature resulting from increases in population. This problem is one of the most important bequeathed by the equilibrium theory to later scientific work.

BIBLIOGRAPHICAL NOTE

A good introduction to the theory of general equilibrium is E. Fossati, *The theory of general static equilibrium*, translated from the Italian by G. L. S. Shackle (Blackwell, 1957). The basic reference is L. Walras, *Eléments d'économie politique pure ou théorie de la richesse sociale* (1874, 4th expanded edition 1900); there is an English translation, based on a collation of the various French editions, by W. Jaffé, *Elements of pure economics* (Allen & Unwin, 1954). A translation of E. Barone's original article, 'Il ministro della produzione nello Stato collettivista' (*Giornale degli Economisti*, 1908), is in *Collectivist economic planning*, edited by F. A. Hayek (Routledge, 1935).

The 'quantity theory' of money comes in two versions: that of Irving Fisher (*The purchasing power of money*, Macmillan, 1911, 2nd edition 1913), and the less rigid Cambridge version (see A. Marshall, *Money, Credit and Commerce*, Macmillan, 1923, and A. C. Pigou, 'The value of legal tender money', *Quarterly Journal of Economics*, 1917) which contemplated variations in the velocity of circulation of money due to factors such as changes in the inter-

est rate. A resuscitation, in a more sophisticated form, of the Cambridge theory has been recently performed by Milton Friedman; see 'The quantity theory of money: a restatement' and his other essays reprinted in *The optimum quantity of money* (Macmillan, 1969).

On the partial equilibrium theories, the basic references are A. Marshall, *Principles of economics* (Macmillan, 1890, 8th edition 1920), and I. Fisher, *The theory of interest* (Kelly & Millman, 1930).

The conceptual difficulties of the theory of capital are well illustrated by Joan Robinson in *The accumulation of capital* (Macmillan, 1956). K. Wicksell's theory is in *Föreläsingar i nationaleconomi* (1901), translated from the Swedish by E. Classen as *Lectures on political economy* (2 volumes, Routledge, 1935).

Chapter 2

ROBBINS' METHODOLOGICAL CONTRIBUTION AND WELFARE ECONOMICS

1. *Robbins' definition of economic science*

In 1932, the British economist Lionel Robbins published a book entitled *Essay on the nature and significance of economic science*, in which he explained the concept of economic activity and, therefore, of economic science, implicit in equilibrium theory.

In order to understand Robbins' scheme properly, it is necessary to begin with the concept of scarcity found in Walras' definition of wealth. As you will remember, Walras defined as 'scarce' that which is at the same time both useful and limited in relation to demand. Now, if wealth is scarce, that is to say, if it cannot satisfy needs up to the point of satiation, it follows that the problem that each individual has to resolve consists in drawing the maximum benefit from the scarce resources available. Equilibrium theory abounds in particular examples of this general problem. The consumer, for example, tends to distribute his income among the various consumer goods in order to maximize his satisfaction (or 'utility'). The producer disposing of certain quantities of productive resources tends to allocate these resources among the various production alternatives open to him in order to maximize his profit. A saver tries to distribute his accumulated savings among various uses in such a way as to maximize future incomes. And so on. It must be repeated that all these are only different aspects of the general problem of obtaining the maximum result with given means, or, which is the same, to obtain a given result with the minimum of means.

This well-consolidated theoretical tradition prompted Robbins to search for a definition of the 'economic fact' which, in his own words, would be not 'classifactory' but 'analytical'. By this Robbins meant a definition that would not select certain facts, certain types of conduct which would be termed 'economic', as distinct from others which would be 'non-economic', but a definition that would indicate in what the truly economic aspect of human conduct

consists. By the standard of such a definition, the adjective 'economic' would apply not to a part or portion of human activity, but to one of its aspects, one of its dimensions. Robbins therefore rejects the 'classificatory' definition, then current, particularly in England, according to which all acts contributing to 'material welfare' would be economic. His criticism is that, even if there were such a thing as a precise concept of material welfare, there would still remain the undoubtedly economic problem of the way in which time and the resources available should be split up among economic and non-economic activities.

In his further investigation of Walras' concept of scarcity, Robbins specifies which conditions are sufficient and necessary for making human action susceptible to economic analysis. There are four such conditions: the first two concern ends, and the other two means. The first condition is that the ends in view must be multiple; the second that the ends must vary in importance and can be ranked in order of priority; the third that the means must be limited; the fourth that the means have alternative uses.

None of these four conditions considered by itself, makes an action 'economic'. The ends may well be multiple, but if the means and the time available are such as to satisfy them all completely, no economic problem arises. The means may be limited, but if they are not susceptible to alternative uses, if they can be used in one way only, their utilization does not present economic aspects. And again, even when the means are scarce and applicable to alternative uses, if the ends are equally important then there is no choice.

What is missing in all these cases is an element essential to the economic aspect of human conduct, and that is the possibility of exercising choice. Only when the means for achieving aims are limited and applicable to alternative uses, and the aims can be ranked in order of priority—only then can human conduct take the form of choice, and so assume an economic dimension. Economics is then a science of human conduct in so far as such conduct takes the form of choice; it is, in Robbins' words, 'the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses'.

The essential character of the scientific work begun with equilibrium theory was made clear by this definition. As Robbins himself recognized, all he did was to organize a set of ideas that were the economists' common inheritance. In fact, the various propositions of economic science are only statements of the various results of the act of choice in particular cases: maximizing results or mini-

mizing means is the general criterion on which choices are based, once the conditions restricting the choice are known. Reading Robbins' *Essay* one becomes aware of the unitary character of economic science, a circumstance well illustrated by the American economist Paul Samuelson who, in 1947, in his book *Foundations of economic analysis*, demonstrated the formal identity of the theories relative to the various areas of economics.

Various consequences follow from Robbins' definition. Here we shall examine those which have had the most importance for scientific work and for economic policy. The questions we shall put are three. The first refers to the deductive character of economic science. The second refers to the so-called 'neutrality between ends' of this science. The third concerns the difference between economics and technology.

2. *Economics as a deductive science*

The concept of economics emerging from Robbins' definition is that of an essentially deductive science, that is, a science that deduces its propositions from certain initial assumptions. These assumptions are, in Robbins' view, certain simple and indisputable facts of experience relating to the way in which the scarcity of goods, which is the subject of economics, is actually revealed in everyday life. If these premises are accepted as corresponding to reality, the generalizations economics achieves by way of deduction also correspond to reality, and describe the general characteristics of human conduct inasmuch as this is economic, whatever the historical and institutional context in which it takes place. On the other hand, the admissibility of the premises on which economic science is built cannot, according to Robbins, be doubted, since these are simple and obvious propositions; it is not even necessary to resort to experiments to check them, as they are the object of our everyday experience. Thus the main premise of the theory of consumer behaviour is that individuals can rank their preferences according to a certain order of priority—an obvious fact needing no psychological theory to support it. Similarly the main premise of production theory is that distinct factors of production exist such that each gives decreasing returns when the others are constant, as is recognized in the classical theory of land rent, where it was pointed out that the application of successive doses of capital to a fixed quantity of land gives rise to smaller and smaller increments

of production. This circumstance, too, is obvious without recourse to special technological justifications.

In this way, economics assumes characteristics similar to those of rational mechanics. As the latter derives its propositions from certain elementary properties of bodies, so economics (which Robbins, in Walras' footsteps, calls 'pure economics') is none other than the development of all the deductions that it is possible to draw from the one initially admitted circumstance: that many ends of diverse importance can be achieved with scarce means applicable to alternative uses. This concept of economics had in fact dominated scientific research for some time, but the rigorous precision which Robbins gave to it greatly contributed to clearing the field of residual uncertainties about the nature of the propositions elaborated by modern economics, and in this way considerably favoured its further progress.

3. *Neutrality between ends*

We now come to the question of the neutrality of economics between ends. It is deduced from Robbins' definition that the economic problem, and therefore scientific reflection on it, arises only where a relationship between means and ends is established. Neither the ends as such, nor the means as such, beyond their reciprocal relationship are of interest for the purposes of economics. In particular, economics is indifferent as to the ends, in the sense that it cannot pass judgement on the ends, for the same reason that it can pass judgement on the most appropriate uses of the means for reaching the ends themselves.

The goals of human action are naturally subject to other types of evaluation—moral, religious political, and so on—but economics cannot do other than take each into account as being merely one of the elements of the economic problem. The specific nature of these goals, however, does not at all concern the formulation of the economic problem, nor, therefore, the reflection upon it. Naturally Robbins specifies, this does not mean that the economist cannot and should not express judgements on the admissibility or the opportunity of any particular ends which a society may propose from time to time—for example in the field of economic policy—but he may not do this as an economist, since his duty lies elsewhere, in indicating how the means at one's disposal can be used to attain those ends in the best possible way. According to this

conception economics is a 'positive science', that is, free from value judgements.

4. *Economics and technology*

Similar considerations, in a certain sense, can be made with regard to means. These too, beyond their relationship with the ends, do not at all concern economics, and this serves to determine a very important difference between economics and technology, the basis of which has only become completely clear as a consequence of Robbins' analysis. The confusion which at first existed on this point derived from the fact that both economics and technology were concerned with the phenomenon of production, and it was not always possible to distinguish in what way one was concerned and in what way the other.

In the light of the concepts put forward by Robbins, the distinction becomes clear. Technology is concerned with the suitability of the means for obtaining certain ends, leaving entirely out of consideration, however, both the degree of scarcity of the means and the order of priority of the ends. Economics, on the other hand, accepting as given what technical knowledge has to say on this point, evaluates the means according to their scarcity and to the relative importance of the ends obtainable by them. The difference can be made clear by looking at particular cases of which the following is an example: automation is a technically better productive process than simple mechanization simply because it can produce more per unit of time, but that is not enough to make it more economical. It could happen that the increase in production obtained by automation satisfies a less important need than others that could be satisfied by using in a different way the additional resources invested in automation.

5. *The problem of the optional configurations of the economic system and the contribution of Pareto*

One of the most outstanding results of contemporary economic thought has been the solution of a problem arising from Robbins' definiton of economic choice. As has been seen, the problem of choice, applied to the behaviour of single economic units, is a

problem that is well defined in terms of maximizing or minimizing a magnitude controlled by an individual. It must now be asked if a similar reasoning to that used in the case of an individual is possible with regard to the whole economic system. In other words, is it possible to define maximizing behaviour for the economic system considered as a whole? A single individual can maximize either his income or his satisfaction ('utility'). But what can the economic system maximize? Is there a magnitude, unequivocally definable, the maximizing of which can be considered as the specific result of a well-functioning economic system? If one accepts the old utilitarian English tradition (still adhered to by Pigou in 1920), the answer to such a question would be relatively simple: the magnitude to maximize in the economic system would be none other than the sum of the utilities of the individuals who make up the system. Now, even without deep and critical reflection, this concept may seem very strange, since one cannot see how the utility or satisfaction of different individuals, even supposing that they are measurable, can be added together.

All the same, it is useful to note that in current economic policy practice comparisons between the utilities of different individuals are continually made. The instituting, for example, of a progressive income tax implies the judgement that the utility loss, consequent to handing over a given sum to the State, is greater for the poor man than for the rich. More generally, most economic policy measures benefit some members of the community to the detriment of others, and if, in spite of this, they are adopted, this means that the advantages are judged to be greater than the disadvantages, which evidently implies a comparison between the utilities of different individuals. But, according to Robbins' approach, comparisons of utility between individuals necessarily bring about value judgements, which are completely beyond the scope of economic science. This does not mean that such valuations should not be made, but that the criteria they are inspired by cannot be drawn from economics. If economics wants to remain a positive science, it must banish all procedures and all ideas which require a comparison between the utilities of different individuals. Our problem therefore cannot be solved, at least not in a strictly analytical sense, by regarding the sum of individual utilities as the magnitude that the economic system should maximize.

Is it then possible, we must ask again, without leaving the field of economics as defined by Robbins, to indicate a maximizing procedure for the system too, so that from the correct solution of this problem of maximization it is possible to derive a strictly economic

judgement on the efficiency of the system itself? It is hardly necessary to note that this question cannot be answered by looking at the value of the national product or national income as the quantity to maximize. In this case it would be a question of comparing heterogeneous aggregates of goods after having made them comparable by means of a certain set of prices. But it so happens that the result of the comparison depends on the prices used: a certain aggregate of goods can be greater, equal to, or less than another aggregate according to the prices adopted in order to carry out the comparison. The quantity to maximize cannot therefore be indicated in any aggregate of goods, because the maximizing operation would not give rise to an unequivocal result.

The solution to the problem has been supplied by contemporary economic thought, taking up an idea expounded by Vilfredo Pareto in his *Manuale di economia politica* in 1906. As has already been mentioned, Pareto is the most important follower of Walras in the field of general economic equilibrium. The most important of his contributions here refers to a question which Walras had already attempted without, however, succeeding in giving it a satisfactory solution. The question is that of the judgement to be given on competition, considered by Walras as the best of all the possible forms of market. Pareto realized that a very important methodological question was implicit in this problem: the problem, in fact, could not be solved without first establishing a criterion on the basis of which one configuration of the economic system could be judged to be superior to another. Only in this way, in fact, could it be determined whether or not the configuration corresponding to a certain market form was the best possible one. The problem of competition therefore immediately gives rise to the problem of what is meant by the optimal configuration of the economic system as a whole.

The criterion supplied by Pareto is the following: a configuration consisting of a set of non-comparable magnitudes is said to be optimum if it is not possible to increase one of these magnitudes without reducing the others. It is important to note, from a logical point of view, the difference between the ordinary concept of optimum and the Pareto optimum (which, incidentally, is known in mathematics as a 'vector maximum' or 'minimum'). Usually, the value of a variable is optimal when all the other values which the variable can assume are inferior to the one under consideration. A configuration is optimal in Pareto's sense when all other possible configurations are either inferior to or cannot be compared with the one under consideration. In the case of general economic equilibrium, this criterion has two specifications. The first refers to

production: given certain amounts of resources and a certain technology, a productive configuration is said to be optimal (or, more precisely, *efficient*) when it is not possible to increase production of a good without reducing production of another good. It is important to bear in mind that there is an infinite number of efficient configurations. The second specification of Pareto's criterion concerns consumption: given certain amounts of goods and given the preferences of the individuals, a certain configuration is optimal when it is not possible to improve the position of an individual without worsening that of another. There is also an infinite number of configurations which are optimal from the consumer's point of view. For the economic system as a whole, the configurations that are optimal from the point of view of both production and consumption are the ones (obviously an infinite number of them) where goods belonging to efficient configurations are optimally distributed among individuals.

It can be demonstrated that the optimal configurations of the economic system are all characterized by the fact that goods are produced, distributed and consumed in such quantities and proportions that the number of units of one good that have to be given up in order to produce an additional unit of another good (*marginal rate of transformation*) is equal to the number of units of the first good that the consumers are psychologically prepared to give up in order to acquire an additional unit of the second good (*marginal rate of substitution*). But since this is, as we have seen in the last chapter, the result reached by a competitive market, it must be concluded that the configuration brought about by competition is optimal. Which optimal configuration is reached by competition depends only on the initial distribution of productive resources. Therefore, unless the government pursues a policy of altering the final distribution of income among the individuals, there is an optimal position for each distribution of resources, and the various optimal positions differ solely in the way in which income is divided among individuals.

Two important consequences follow. Firstly, the optimality criterion defined by Pareto allows Walras' argument, that competition is the best possible market form, to be proved. Secondly, it seems obvious that Pareto's criterion cannot be used to compare two situations, both optimal, which differ only in the distribution of income. Whether a certain distribution of income is better than others is a problem left open by Pareto.

6. *Welfare economics*

Since 1934, a vast amount of literature relating to optimal positions in the economic system has developed on the basis supplied by Pareto. This literature is opposed to the utilitarian approach which, culminated in 1920 with the book *The economics of welfare* by the English economist Pigou. The new doctrine defined itself as the 'new welfare economics'. Here the principal names are: Hicks, Lerner, Kaldor and Little in Britain; Hotelling, Bergson, Lange, Samuelson, Arrow and Debreu in the United States; Maurice Allais in France. The problems posed by this literature are numerous, and it is therefore not possible to examine them all. Two only will be dealt with here.

The first, suggested by some authors, concerns an extension of the Paretian criterion. According to these authors, in order to be able to say that configuration A is better than B, it is not necessary to exclude the possibility that, in moving from B to A, someone is worse off. It suffices that, once the second configuration is reached, those who are better off have additional resources sufficient to compensate those who have come off worse. According to this conception, it is not necessary for compensation to be actually paid; it is sufficient that this is possible in principle. Thus, for example, the repeal of the Corn Laws in England should be judged as a policy that improved the general welfare of the country in that the gain made by the consumers was such that it would have allowed them to indemnify the farmers hit by this measure. This extension of Pareto's criterion is obviously important because it makes it possible to express an objective rather than a value judgement on a much larger number of government policies.

The second problem is rather subtle and, at first sight, of importance only from a purely theoretical point of view. All the same, this too has many practical implications. We have already noted that the Paretian criterion can be used to prove that the competitive configuration is optimal. Modern welfare economics, on the other hand, has shown that the converse is also true; that is, that any optimal configuration can be conceived only as a competitive configuration. If you think about it, it is this second proposition which is of more practical importance. It is, in fact, possible to deduce from it that only competition gives rise to optimal configurations, and that, therefore, all other market forms (monopoly, oligopoly, imperfect competition, etc.) necessarily give rise to non-optimal situations; that is, to situations that can be improved in the sense that one could move to other configurations in which some would

be better off and none would be worse off. This laid the theoretical foundations of anti-monopoly legislation.

However, this one-to-one correspondence between optimal configurations and competitive situations also has a further important consequence for economic planning. In order to see this, one must note that the said correspondence throws new light on the concept of price. We have already seen, with reference to equilibrium theory, that it is possible to have a concept of price which is, so to speak, generalized, in the sense that it does not necessarily imply a market transaction. On the basis of the findings of the new welfare economics, we can now add that the equilibrium prices, being the only ones corresponding to an optimal situation, not only *can* but *must* serve as a guide to the rational economic planner. In other words, in order to plan rationally, it is necessary to make all calculations on the basis of a set of prices that satisfy simultaneously the subjective and objective conditions of general equilibrium. Only in this way is the best use of available resources guaranteed. We will come back to this in Chapter 10. For the time being, we shall simply add that, from what has been said, it follows that such a plan must fix the distribution criteria if the problem of the planner is to be determinate; that is, if we are to avoid ending up with an infinity of solutions. But, a word of warning: the limitations of equilibrium theory—in particular the fact that it ignores the problem of growth—mean that it needs susbtantial integration as a foundation for economic planning, without, however, losing for that its relevance and validity.

BIBLIOGRAPHICAL NOTE

The basic references on the scope and method of economics are L. Robbins, *Essay on the nature and significance of economic science* (Macmillan, 1932), and P. A. Samuelson, *Foundations of economic analysis* (Harvard University Press, 1947). V. Pareto's original work is in Italian (*Manuale di Economia Politica*, 1906), but was later expanded and improved by the author, particularly in the mathematical appendices, in the French version *(Manuel d'économie politique*, M. Giard, 1909, 2nd edition 1927).

On traditional welfare economics see A. C. Pigou, *The economics of welfare* (*Macmillan*, 1920, 4th edition 1932). On the 'new' welfare economics, see A. Bergson, 'A reformulation of certain aspects of welfare economics' (*Quarterly Journal of*

Economics, 1938), I. M. D. Little, 'The foundations of welfare economics' (*Oxford Economic Papers*, 1949), and Samuelson's *Foundations*. A good review, with a vast bibliography, is E. J. Mishan, 'A survey of welfare economics, 1939–59' (*Economic Journal*, 1960). For a good textbook treatment, see J. V. de Graaff, *Theoretical welfare economics* (Cambridge University Press, 1957).

Chapter 3

SCHUMPETER AND THE THEORY OF ECONOMIC DEVELOPMENT

1. *Economic growth and economic development*

The theory of general economic equilibrium left open the essential question of economic development, for it must be remembered that this theory assumes that technology and consumer preferences are unchanging. Under such conditions, the economy can only tend towards a situation in which output is either stationary or, if the population is increasing, grows at the same rate as the labour force, without any change in quality or method of production. This purely quantitative pattern of economic expansion was to be the subject of the first crop of growth models, produced in the 1940s and 1950s (see Chapter 9 below). Here we are concerned with economic development, by which we mean expansion of the economy accompanied by structural changes in the productive system, and by changes in the quality and composition of the final product.

In 1912, the Austrian economist Joseph Schumpeter published a fundamental book on the subject, entitled *The Theory of Economic Development*. Schumpeter went to the United States in 1932 and there made other distinguished contributions to the theory of economic dynamics, the most important of which are to be found in his *Business Cycles* (1939), and *Capitalism, Socialism and Democracy* (1942). Basically, Schumpeter's ideas changed very little after 1912. His subsequent works develop and discuss particular problems, without changing his theory's general design.

Schumpeter takes Walras' equilibrium as a starting-point for his theory of development. He considers, on the one hand, that Walras' system is indispensable for highlighting the fundamental relationships of an economic system, and, on the other hand, that it is impossible to understand the development process without first explaining how such a process is sparked off by breaking the stationary equilibrium. One gathers from the many references in his books that Schumpeter considered Walras to be the greatest of all economists, which is an indication of Walras' key position in the history of modern economic thought.

2. *Innovations and entrepreneurial activity*

The singularity of Walras' general equilibrium theory is that it portrays a process continually repeating itself in the fields of both production and of consumption. Once competition has pushed the system into a position of maximum efficiency, this configuration repeats itself over and over again. The notable consequence for the business world is that the management of a firm is reduced to mere routine. Each firm goes on producing the same types and the same quantities of goods, combining the productive factors always in the same way.

In Schumpeter's view, this stationary state is broken within the ambit of production, as a result of often radical changes in methods of production which start off the development process. Such changes are classified by Schumpeter as follows. Firstly, the introduction of a new good—one that is unfamiliar to consumers—or a new type of a given good. Secondly, the introduction of a new method of production, or of a method known but yet untried in the branch of industry into which it is being introduced; this new method of production is not necessarily the result of a scientific discovery, but may perhaps merely consist of a new way of presenting a commodity. Thirdly, the opening up of a new market, that is, a market which is new for a given industry, in the sense that the products of this industry had never had access to it before, whether that market already existed or not. Fourthly, the opening up of a new source of raw materials or of semi-finished goods, again whether this source existed before or was newly created. Lastly, the reorganization of an industry, as by creating a monopoly, or by breaking a monopolistic situation. These changes are collectively referred to as 'innovations'. Schumpeter also defines the 'entrepreneurial act' as the introduction of an innovation into the economic system, and 'entrepreneur' as one who carries out this act. Entrepreneurs and entrepreneurial activities are therefore phenomena specifically pertaining to development and are completely unknown in the stationary state, where the management of the business firm is mere routine, indistinguishable from any other type of work.

The distinction between entrepreneur and manager is therefore fundamental, even if the two are often one and the same person or collective body. The decisive difference between the two functions can be seen by reflecting on the difference, for example, between the decision to introduce a new method of producing woollen cloth and the decision to buy the amount of wool required by the current

method of production.

3. *Profit*

Profit is the result of entrepreneurial activity. The most obvious example of an innovation generating a profit is that of the production of a good already in use at a lower unit cost than that of any other firm, because it is produced by a new method requiring less of one or more factors per unit of output. In this case the entrepreneur buys the productive factors he needs at the current prices. The level of these prices results from the conditions in which the 'old firms' operate, and the entrepreneur sells his product at the current price—which also corresponds to the costs of the 'old firms'. It follows that his receipts exceed his costs, and the surplus is what we call profit.

Two other types of innovations—the reorganization of production and the finding of a new and less expensive source of productive factors—give rise to profits in precisely the same way as described. When new goods are created which better satisfy needs that are already catered for by existing products, the opportunity of making a profit arises from the fact that the higher price charged for the improved product often exceeds its cost, even if this cost is higher than that of the old product. In the case where the innovation consists of finding a new market, where a certain commodity, even if already produced and sold elsewhere, was formerly unknown, a profit can be made because the new buyers are prepared to pay high prices which have no connection whatever with costs. Similarly, in the case of the creation of new goods which satisfy new or formerly unsatisfied needs, the prices, at least initially, bear no relation to the costs.

As innovations do not exist in the stationary state, profit cannot exist either. Here, too, we have a phenomenon typical of development, but which, in the course of development, continually tends to disappear even if the conditions for its existence keep reappearing. Once, in fact, that profit has been generated in one point of the system, the condition which generated it—that is, the innovation —becomes generalized, and the process of competition tending to bring prices to the level of costs will determine the disappearance of the profit itself. It should be understood that it disappears only from the firms' point of view, because in reality it does not disappear at all, but has spread throughout the entire economic

system giving rise to an increase in welfare with no additional effort spent in production. But even from the firms' point of view, if the process of introducing innovations does not stop, profit keeps reappearing so that at any given moment business revenues having the nature of a profit are found. It is naturally possible, and in fact often happens, that the mechanism of competition does not work perfectly. In this case the profit, or a part of it, does not spread through the entire system and tends to remain within the ambit of the firm. It then ceases to be a profit, because it is no longer the result of innovations, but is maintained within the ambit of routine activity and becomes a monopoly rent.

4. *The stages of capitalistic development*

Having defined the figure of the entrepreneur in general terms, the question arises of who actually performs the function of entrepreneur in a capitalistic economy, which, after all, is the kind of economic organization specifically studied by Schumpeter. There is then the question of whom the profit goes to in such an economic organization—a question related but not identical to the first, for although the entrepreneurs' presence is necessary for generating a profit, this does not mean that he is necessarily the one to benefit from it.

To answer these questions one has to consider that, in Schumpeter's opinion, an innovation generally means building new plants or at least radically transforming old ones. Although this is not always necessarily so, it can be assumed that all innovations not falling into this category are minor ones, that is, not characteristic of the development process. Besides, the creation of new plants may be the result of new firms being formed, or of existing firms being expanded. Here Schumpeter distinguishes two stages in the history of capitalism; the first he calls 'competitive capitalism' and the second 'trustified capitalism'. The first stage is characterized by firms that are not too large in relation to the over-all size of the market; here, the introduction of innovations generally means creating new firms. In the second stage the converse is true: large firms tend to predominate, firms which are in a position to keep the innovatory process within themselves, so that innovation does not imply the appearance of new firms in competition with the old ones. Having said this, and keeping in mind that it is never easy to identify who is the entrepreneur, because nobody is exclusively

and at all times an entrepreneur. Schumpeter remarks that in the age of competitive capitalism the function of the entrepreneur was generally carried out by the owners of the firms themselves. But the question is much more complex in an age in which large combines predominate where the entrepreneurial function may be carried out by the person in control of the firm—that is, by the majority shareholder in a corporation—or by the directors, or even by ordinary managers, and it may lie with the individuals or with collective bodies.

Once a profit has been generated, whether it goes to the entrepreneur or not is an institutional matter. In the case of family firms, the profit normally goes to those who have actually carried out the entrepreneurial activity, and forms the origin of those great fortunes upon which industrial dynasties are founded. On the other hand, in an industrial system based on large corporations, profit, as such, goes to the firm, and its distribution then becomes a problem of internal policy. Those who benefit may be the shareholders, or the members of the board of directors or even the managers and workers, quite independently of who in fact acted as entrepreneur. Although this problem of distribution is indeterminate, it nevertheless remains well established that in Schumpeter profit cannot, in any circumstance, be a reward for risk, as other economists have often declared. Schumpeter points out that the risk is borne by by the capitalist and not by the entrepreneur, for the entrepreneur bears the risk only in so far as he is also an owner of capital.

5. *Credit and saving*

If it is accepted that innovations, or at least those which are really important for economic development, are embodied in new plants, the problem immediately arises of how these innovations are to be financed. In the stationary state, each firm finances its operations out of current receipts. But the entrepreneur who is to build a plant where his innovation is to be implemented needs new buying power to acquire certain productive resources. These resources are thus withdrawn from old uses and directed to new ones required by the innovation. These new means of payment are made available by credit, which for Schumpeter is another of the fundamental characteristics of economic development. In a planned economy, implementing an innovation would require an order from the planning authority to divert productive resources from their present uses to

the new one. In the same way, in a capitalistic economy, credit, in the hands of the entrepreneur, serves a similar function, because it allows the entrepreneur to use part of the resources of the system for his own purposes.

Naturally one imagines that an innovation is financed not by credit but by saving, that is, by funds put aside by households or by firms out of their current income. But, according to Schumpeter's logic, this possibility has to be excluded, since in the stationary state saving either does not exist or exists only in minimal quantities. The main source of saving, in fact, is none other than the profits subtracted from the process or diffusion which is determined by competition, and it is only in a growing economy that innovatory investment can be financed without recourse to credit. Schumpeter does not say that this logical course necessarily has to correspond to the historical course. That is, if the creation of credit by the banking system is logically the beginning of the development process, this does not mean that it must also precede it historically. What actually happened is that, at the beginning of capitalistic development, firms were small enough to be financed with savings accumulated in preceding economic systems.

It could perhaps be inferred that, in Schumpeter's view, this initial stage of 'primitive financing' was followed by two other stages of capitalism, as his distinction between 'competitive capitalism' and 'trustified capitalism' seems to suggest. Corresponding to the first stage would be the great development of the banking system, which showed itself through the performance of its *essential* function of financing innovations. At the same time, the banks were also financing current business transactions, and, on the other hand, some innovations were being financed by dynamic profits. But the basic characteristic of this stage would remain the fact that the banks continually re-created the financial conditions for development, which were systematically destroyed by competition. In the second stage, some characteristics of the initial period are somehow reproduced, even if at a different level and in very different conditions. The emergence and consolidation of firms of ever-increasing size, and the diffusion of practices aimed at obstructing competition are, in fact, phenomena tending to create permanent sources of finance within the firms themselves, and, therefore, to confine the banks to their secondary function of financing current transactions.

Schumpeter's distinction between 'competitve capitalism' and 'trustified capitalism' throws light on another matter of great importance for economic theory, that is, the definition of the concepts of competition and monopoly. We have already mentioned that in equilibrium theory competition is defined in purely static terms. A competitive market is one that consists of a large number of firms, all producing exactly the same good, and so small compared with the size of the market as to have no influence, individually, on the price. Obviously, such a situation is not easily found in real life; we shall see in the next chapter how the traditional theory of competition can be criticized, even from a purely static viewpoint. But what we are interested in, here, is Schumpeter's criticism of the Walrasian-Marshallian idea of competition—a critique that differs from the others to be examined because it is put from a dynamic point of view, in keeping with the general characteristics of Schumpeter's system.

According to Schumpeter, in a capitalist economy the real competition is not between small firms producing the same commodity, but between the innovating firms—those where entrepreneurial activity is present—and the others. The competition is not between identical goods, produced in the same way, but between the new products and the old, or between the new methods of production and the antiquated ones. This competitive process has also been called the process of 'creative destruction' by Schumpeter, a name emphasizing that competition is the effect of innovations on existing firms.

This conception of competition brings with it a conception of monopoly which is also different from the traditional one. The first thing to note is that the introduction of innovations inevitably brings with it a certain degree of monopoly. Before it becomes widespread, the innovation is, in fact, the monopoly of the entrepreneur, and the profit he derives from it is due to this monopoly. Of course, this is not an absolute monopoly, like the one envisaged by the traditional theory, but only a temporary monopoly. In normal circumstances, such a monopoly would disappear in the dynamic process of competition, which at first diffuses the innovation that made the monopoly possible, and then subjects the innovation to comparison with further innovations. But it is also possible that some firms manage to avoid this competitive process. In this case, their profit would gradually assume the characteristics of monopoly rent, and their position in the market would gradually

become one of permanent monopoly. This has happened with increasing frequency in the history of capitalism. All the same, Schumpeter warns against considering so-called 'monopolistic practices' as symptoms of a pathological situation which would slow down the rate of development. Here, his arguments are twofold. Firstly, a large firm where monopolistic practices are more common, far from being an obstacle to innovation is very often where they are most likely to occur, because research and scientific experimentation are easier in large productive units. Secondly, suspending the competitive process, even for a long period, can be viewed as a kind of insurance against the risk arising from adopting far-reaching innovations. Such a risk might be unbearable in rapidly changing market conditions, and to offset it one would have to aim at stabilizing the market by patents, industrial secrets long-term contracts, cartel agreements, rigid prices, etc.

Passing from competitive capitalism to trustified capitalism—that is, from the stage in which innovations are prevalently embodied in new firms to the stage in which they are carried out prevalently by existing firms—does not involve, according to Schumpeter, any slowing down of the pace of economic development, nor a drop in quality; it is even possible for development to be accelerated by the change. Schumpeter, therefore, rejects the thesis, advanced from many quarters, that capitalism is doomed to end in a final crisis because of its inherent economic mechanism. He is, however, convinced that capitalism, to use his own words, 'cannot survive', but bases his conviction on non-economic considerations. Schumpeter's ideas on this subject will be examined in the next chapter, together with other theories on the final destiny of capitalism.

7. *The trade cycle*

According to Schumpeter, the development of the capitalist economy which is generated by innovatory processes is not continuous and uniform, but a sequence of cycles. The trade cycle is not a secondary aspect of capitalism, but, as Marx had already emphasized, is the very way this type of economy grows. For Schumpeter, the cause of the cyclical movement is that innovations are not uniformly distributed through time, but tend to be concentrated or, as he puts it, to 'crowd together' at certain periods. Schumpeter regards this as an historical fact which he explains by noting that,

in order to introduce an innovation, a series of 'social resistances' must be broken. These resistances, which are opposed to anything substantially new, which breaks with the traditional patterns of production, are easier to overcome each time someone more enterprising than usual leads the struggle against custom. On the other hand, this concentration of innovations is, sooner or later, bound to peter out, since, as time goes on, the mass of new goods resulting from the innovations puts increasing pressure on the market. This, together with the increasing volume of debt repayments by the entrepreneurs, has a dampening effect that lowers profit expectations and slows down the introduction of innovations. This mechanism is for Schumpeter the fundamental factor determining fluctuations. Note, however, that the uneven distribution of innovations over time seems more like an effect than a cause of the trade cycle and that, if anything, it tends to amplify the fluctuations. Schumpeter's contribution, therefore, though considerable, needs to be integrated into a more general theory of the trade cycle such as the one developed in modern times after Keynes. We will come back to this later.

According to Schumpeter, the lag between the time when an innovation is adopted and the time it begins to bear fruit in the form of goods to be sold varies according to the innovation itself; this is the origin of cycles of different frequencies. Schumpeter, on the basis of empirical observation, distinguishes three kinds of cycles. The first are the so-called *long waves* (*Kondratieff cycles*), lasting between fifty-four and sixty years: of those recorded, the first lasted from 1783 to 1842, the second from 1842 to 1897, and the third has not yet finished. The second type of cycle lasts between nine and ten years (*Juglar cycles*), so that the Kondratieff cycle contains about six of them. The third type is cycles lasting about forty months (*Kitchin cycles*), so that a Juglar cycle contains about three of them. Each of the historically recorded cycles is associated with a certain group of innovations.

The cyclical nature of the 'long waves' is a controversial question since, according to other students of the trade cycle, the very long-term movements are associated with factors outside the economic mechanism in its narrow sense. Even if this thesis, and not Schumpeter's, were found to be true, it does not affect his central idea that capitalistic expansion is necessarily cyclical in nature. The problem of modifying this proposition in the light of more recent developments in the capitalist economy can be solved, within the Schumpeterian framework, only by taking into account extra-economic changes that have taken place.

J. A. Schumpeter's basic ideas are in *Theorie der wirtschaftlichen Entwicklung* (1926), translated by R. Opie as *The theory of economic development* (Harvard University Press, 1934). The stages of capitalism are in *Business Cycles* (2 volumes, McGraw-Hill, 1939), where economic fluctuations are examined from the historical, statistical and theoretical points of view. The competition between old and new commodities and the future of capitalism are discussed in *Capitalism, Socialism and Democracy* (Allen & Unwin, 1942, 4th edition 1954). For a concise presentation of Schumpeter's theory of economic fluctuations, see R. M. Goodwin, 'Innovations and the irregularity of economic cycles' (*Review of Economics and Statistics*, 1946).

Schumpeter is also the author of a series of bibliographical essays, reprinted in the volume *Ten great economists: from Marx to Keynes* (Allen & Unwin, 1952), and of a massive *History of economic analysis* (3 volumes, Oxford University Press, 1954), published posthumously which is probably one of the best ever written.

Chapter 4

THE NEW MARKET THEORIES

1. *Sraffa's criticism of the theory of perfect competition*

Equilibrium theory had concentrated on the study of two extreme market forms: perfect competition and monopoly. Between 1926 and 1933 economic theory, particularly in Anglo-Saxon countries, subjected to penetrating criticism the claims that these two extremes were representative of capitalist markets in the real world. This criticism developed along lines which were rather different from those advanced by Schumpeter. While Schumpeter based his analysis on a dynamic model assuming all technological factors variable, and indeed arguing that such changes were the origin of the competitive process, the new theories remained within the field of static equilibrium theory and sought to show that, even assuming no technological change and constant demand conditions, neither perfect competition nor monopoly coud be taken to represent reality.

This new approach stems from an article by the Italian economist Piero Sraffa, published in English in 1926: *The laws of returns under competitive conditions*. Sraffa was critical of the proposition that the perfectly competitive firm limits its expansion of output due to rising costs which occur after a certain level of production has been reached. According to the traditional theory, the price is taken as given by the competitive firm and thus, if the average cost increases after a certain volume of output has been reached, there will be a point beyond which further expansion of output will reduce profits. Nevertheless, according to Sraffa, experience shows that what limits expansion of output in competitive firms is not the behaviour of costs but instead demand conditions. More precisely, every attempt to increase production requires a reduction in selling price. If this is true, we must conclude that, for the firm which faces competition, price is not a constant but a decreasing function of the volume of sales, and that because of this the firm finds itself faced not with a unique price but with an entire demand schedule.

We must now consider what market structure lies behind the

assumption that price is given for the firm. As we have already said, it was assumed that each firm is very small in relation to the market as a whole. But, in order to talk of a market, it is necessary to assume that consumers of a given commodity are completely indifferent as to which firm they purchase this commodity from. In this way, each firm is assumed to be part of a vast market for perfectly homogeneous products, in which it is not possible for any firm to sell its product at a higher price than that of any other firm because in this case it would lose all its customers. This results in a unique price which is thus taken as given by each individual bidder.

If, however, one assumes instead that for any given firm price is a decreasing function of sales, the image of this homogeneous market fades and one must suppose that each firm possesses some sort of market peculiar to itself. This in its turn implies that purchasers will not be indifferent as to which firm they buy from. This lack of indifference may be due to various reasons, among which Sraffa enumerated long-standing custom, personal acquaintance, confidence in the quality of the product, proximity, knowledge of particular requirements, the possibility of obtaining credit, the reputation of a trademark, special features of modelling or design in the product which are developed principally to distinguish it from the products of other firms.

It is clear that the existence of as many individual markets as there are firms results in a close analogy between this market structure and that of pure monopoly. In the case of monopoly, price is also a decreasing function of sales and the only difference is that the product of the competitive firm possesses a greater degree of substitutability with the products of other firms. The result of this important difference is that when such a firm raises the price of its product, the decrease in quantity sold will be greater than would be the case for a monopoly because in the first case some of the firm's clients may purchase substitute goods, a course of action excluded by definition in the case of monopoly. In technical language, we say that under competition the elasticity of demand is greater than under monopoly.

The competition that we are discussing is thus rather different from perfect competition. The element which synthesizes the various aspects of this diversity is the circumstance that, while a unique price exists under conditions of perfect competition for the goods produced by a myriad of firms, in the type of competition discussed by Sraffa every firm can sell at its own price, even if (and in this it differs from monopoly) such prices are not independent of one another.

The consequences of this criticism of equilibrium theory by Sraffa are rather destructive. Even though the criticism was directed mainly against Marshallian theory, the gravest consequences are those which refer to general equilibrium theory, since the abandoning of the hypothesis that competition is perfect renders impossible the development of a theory of general equilibrium. Moreover, even for Marshallian partial equilibrium theory, the abandoning of perfect competition or of monopoly means the risk of reducing economic discussion to the listing of an inexhaustible series of particular cases.

Nevertheless, several years after Sraffa's article, efforts were made in England and America to reconstruct a market theory which, while retaining the degree of coherence and of rigour which are the properties of the perfect competition and monopoly theory, would also integrate the propositions advanced by Sraffa. This attempt is expressed especially in two books which appeared simultaneously in 1933: *The Economics of Imperfect Competition* by Joan Robinson in England, and *The Theory of Monopolistic Competition* by Edward Chamberlin in America. To be precise, while Mrs Robinson's work is directly related to Sraffa's position, that of Chamberlin is developed along independent lines. Nevertheless, as we shall see the two books refer essentially to the same subject.

At this point it is convenient to clear up several questions of terminology which for some time have confused the discussion of these market structure problems. We must now see whether the description of 'imperfect competition' adopted by Joan Robinson, and that of 'monopolistic competition' adopted by Chamberlin, cover the same or different concepts.

Let us start by examining Joan Robinson's treatment: she notes that the perfectly competitive market is based on the assumption that buyers will all react in the same way to the differences in prices charged by various producers. However, in real markets the buyer takes into account many factors other than the price. Leaving aside inertia or ignorance, which restricts instantaneous movement from one vendor to another as soon as a difference appears in the market prices, we can suggest many good reasons why a buyer may prefer one vendor to another. Mrs Robinson gives the following examples: the location of the vendor, transportation costs, the guarantee of quality provided by a well-known name, different facilities offered by various producers, the generosity of sales services, publicity, etc. All these reasons destroy the homogeneity of the perfectly competi-

tive market, and introduce a market imperfection which is manifested in the fact that each firm commands the loyalty of a certain group of customers who cannot be dissuaded from buying from that firm by simple price considerations.

The concept on which Chamberlin's construction is based is one of product differentiation rather than of market imperfection. What he means by this is that goods produced by competing firms are not in fact identical due to differences in their nature or are not considered identical by purchasers for various reasons. Chamberlin gives the following examples of differentiation: the existence of a brand name, peculiarities in the quality, form or packaging of the product, the vendors' location, the efficiency reputation or sales service of the vendor, etc.

It can be easily seen that these concepts of market imperfection and of product differentiation both refer to the same real world problem. Regardless of what name one attaches to this phenomenon, the distinguishing characteristic is that various reasons exist as a result of which the goods of the firms are not considered to be identical by the various purchasers. The causes which give rise to this in practice are infinite and one can give many examples. Goods produced by different firms may be competitive but not identical to buyers because there is a real difference in quality between the products (between different makes of radio, cars, bicycles, refrigerators, shirts, socks, coffee, etc.); or because, despite the fact that real differences are negligible, buyers have some reason for believing that differences do exist; or because the conditions under which products are sold vary (for example, the way in which they are packaged, the different credit facilities which may be available or, even, the greater or lesser degree of courtesy of the shop assistants which induces purchasers to patronize one shop rather than another); the different spatial distribution of the shops, the varying reputations of producers and sellers, and so on are other possible reasons.

In the following discussion we shall follow Chamberlin's analysis, for he is more concerned than Joan Robinson with describing a market structure which lies between perfect competition and monopoly.

3. *The theory of monopolistic competition*

It is clearly not possible in this book to elaborate in detail all the

technical aspects of Chamberlin's analysis. We shall limit ourselves, therefore, to illustrating several essential concepts which will subsequently permit us to undertake a critical consideration of the theory of monopolistic competition. First of all, let us recall the essence of this concept. What we are discussing is a form of market in between that of perfect competition and monopoly in the sense that the group of firms are competing because, as distinct from what occurs under monopoly, the demand for the goods of each firm is not unaffected by the existence of the other firms; thus the presence of these other firms is a fact which each individual firm cannot ignore. On the other hand this type of competition contains an element of monopoly because, as distinct from what occurs under perfect competition, each firm has its own particular product which, if not different from that of other competing firms, is nevertheless quite distinct.

It follows from this that the type of competition to which the firms are subjected is not only price competition, as in perfectly competitive markets, but is also product competition. In other words, the various producers can compete with one another by modifying the quality of their product in such a way as to attract clients away from other firms. In perfectly competitive markets competition based on product quality variation is not possible, as the homogeneity of the products of the various firms is part of the definition of perfect competition itself. Competition based on quality can occur in two ways. The sellers can fix the quality of the product leaving the purchasers to be guided only by their own preferences and by their own tastes in the choice of one firm's product rather than that of another. Alternatively, the vendors can advertise the quality of their own products so that buyers are influenced not only by their own preferences, but also by deliberate actions of the vendors. Thus, in the first case, those who sell accept the demand conditions as they are freely expressed in the market, while in the second case they seek to modify (often radically) such demand conditions.

Thus the examination of monopolistic competition is much richer than that of perfect competition. Chamberlin, nevertheless, sought to conduct his examination along lines analogous to those adopted by Marshall for perfect competition. Marshall distinguished two phases in the study of competitive equilibrium. The first phase referred to the equilibrium of the individual firm; the second phase referred to the equilibrium of the industry, which comprises all the firms producing the same product. The conditions for the equilibrium of the firm derive from the maximization of

profits, while the conditions for industry equilibrium result from the situation that the industry under consideration does not offer opportunities for earning bigger profits than those existing in other industries. It is for this reason that, in addition to the other conditions which characterize perfect competition, so-called 'freedom of entry' is assumed or, in other words, there is a complete absence of barriers to the influx of firms to any given industry. The analogy with Marshallian treatment can be maintained for monopolistic competition in so far as it is possible to find a substitute for the concept of an industry. It is obvious that the Marshallian industry cannot exist in monopolistic competition since in this form of market, by definition, no two firms exist which produce identical goods. Chamberlin thus substituted for the concept of industry that of the 'group', meaning those various firms whose products, while not identical, are nevertheless to a considerable degree substitutes. Naturally, the group notion is rather vague compared with that of industry: between Marshallian industries there are clear boundaries separating one from another, but it is not possible to state exactly where one group finishes and another begins. All the same, Chamberlin holds that, not withstanding this indeterminacy, the concept is sufficient to permit analysis.

Such analysis, then, as in Marshall, occurs in two steps, referring respectively to the equilibrium of the firm and the equilibrium of the group. In addition, both steps may be examined under the assumption that the vendors accept demand conditions as given, or that they seek to modify these through advertising. The equilibrium of the firm is determined in a similar fashion to that of monopoly, given that for the single firm the most important factor is the monopoly element and thus price is a decreasing function of quantity sold. Regarding group equilibrium, the theory succeeds in defining an equilibrium position at which, due to freedom of entry into the group under consideration, the price and the unit cost coincide for each individual firm. Thus we arrive at a position analogous to the competitive industry equilibrium, the only difference being that, while under perfect competition industry equilibrium results in every firm producing that quantity which corresponds to mimimum unit cost, in a group equilibrium of monopolistic competition the firms produce a smaller quantity than under perfect competition, and the average cost is not at the minimum. If, as is usual, one defines full productive capacity as that which corresponds to the minimum average cost, one must conclude that under monopolistic competition the group equilibrium results in underutilization of productive capacity.

This theory may be criticized in many ways. We will limit ourselves here to setting out three criticisms from those which seem most relevant.

The first criticism concerns not only the theory of monopolistic competition, but also the theory of pure monopoly. As we have already noted, when we abandon the hypothesis that for the firm price is given, we must recognize that the firm itself faces an entire demand schedule. On the other hand, the assumption that a firm knows the demand schedule for its products—that is, that it can recognize, without experimenting, what quantitiy will be sold for every possible price—is not plausible. Various empirical studies of large samples of firms have confirmed that this is not possible. It is clear, on the other hand, that if the demand curve cannot be drawn up, then the whole analysis of monopolistic competition, and indeed all other forms of market which are not perfectly competitive, faces an insoluble difficulty.

The second criticism also concerns the equilibrium of the firm and refers to a variable peculiar to monopolistic competition, that is to say, the quality of the product. Quality is not something which is directly measurable and thus its treatment is necessarily different from that of price and quantity sold. So far, despite various efforts that the theory of monopolistic competition has made to include quality as a variable in the technical analysis, it has not succeeded other than to explain, in a complicated fashion, the tautological proposition according to which a firm seeking to maximize profits will choose that quality which will be most profitable.

The third criticism refers to the equilibrium of the group. The proposition that it is possible to define an equilibrium for the group, analogous to that of the perfectly competitive industry, is valid only if we make the rather restrictive assumption that the firms which form the group have identical cost curves. Without this assumption, the process of entry of new firms cannot result in the simultaneous attainment by each firm in the group of a situation in which price and average cost are equal. Thus, each firm will tend towards such a position, but some will reach it before the others and, if the process of entry continues, the firms which reach it first will find themselves making losses and therefore will tend to leave the group. Thus, along with a process of entry, there will also be a process of exit and there is no reason to suppose that these two processes will cancel each other out. The resulting situation is therefore permanently unstable. The very nature of monopolistic competition obliges us to exclude the possibility that firms have the same cost functions. It is thus impossible for the theory to define a position of

equilibrium for this form of market. How serious this is from the point of view of the theory is clearly seen from the fact that the analysis of the group is essential for the theory of monopolistic competition which, in the absence of such an analysis, would not be distinguishable from the theory of monopoly.

The uncertainties and the relative doubts regarding the firm's equilibrium and the impossibility of defining an equilibrium for the group make the theory of monopolistic competition an instrument of analysis of rather dubious theoretical validity. Certainly it has the merit of wishing to base its propositions and assumptions more closely on reality than does the theory of perfect competition, but any progress which it has made in the direction of greater realism has been accompanied by a diminution of rigour in the theory.

4. *The problems of oligopoly*

The same spirit of criticism which resulted in the theory of monopolistic competition at the same time generated a return to research into other market forms which, while being recognized for some time, had not received much attention after the dissemination of the theory of equilibrium based on perfect competition. The studies of oligopoly are of particular interest in this respect. Oligopoly, as the name indicates, is that form of market in which the firms in competition are few. This term 'few' has a precise significance in economic theory which can be clarified in the following way. Consider a firm in a monopoly position or a firm belonging to a group in monopolistic competition. Such a firm knows, by definition, what the reactions of the market will be to its actions; that is to say it can tell what quantity it can sell at any particular price, or indeed what price it should charge for any particular quantity it wants to sell. This means, in other words, that such a market has a structure of its own, independent of the actions of the firm under consideration, a structure which is reflected in a well-defined demand schedule. Here we are of course referring to the simple possibility of defining a demand schedule for the firm, whether the firm knows this schedule or not.

In the case of monopoly this arises because there are no other firms in the market and therefore the monopolist's demand schedule coincides with the demand schedule of the entire market. In the case of monopolistic competition there are many other firms in the

group. and each one, considered individually, is influenced to a very slight degree by the actions of the firm in question. What happens is that, if one firm raises its price, it loses customers to the other firms, but these lost customers distribute themselves among the numerous other firms in the group; thus none of these is noticeably influenced when considered separately from the others. From this we see that any individual firm possesses a well-defined portion of the aggregate market which permits it to predict the reactions of its own market to its actions.

But if we look at a situation in between these two extreme market forms and suppose that the number of competing firms is sufficiently small—no matter whether this competition is perfect or less than perfect—then the conduct of a firm will noticeably affect all the other firms *individually*. These, in turn, will react in a certain way which will modify the portion of the market which belongs to the first firm. This firm, therefore, in determining what action it should take, must take into account the reactions of all the other competing firms, but these other firms in their turn seek to make their plans on the basis of the actions of the first firm. Thus we have the same situation that would occur in a card game in which each player, before deciding his own move, takes into account the reactions of the other players. With traditional analytical instruments, using cost and demand curves, it is not possible to show a determinate solution for the oligopoly problem. Modern criticism has in fact shown that all the solutions put forward by preceding economists, and in particular those of Cournot and of Edgeworth, are founded on special assumptions which cannot be considered as precise and exhaustive descriptions of reality.

Similar conclusions were reached in the case of a market form which also underwent considerable study in this period, that of bilateral monopoly, which is a situation where a good has one supplier and one buyer. This situation is not unusual in the productive factor markets, as might occur, for example, if in a certain region a large industry is the only purchaser of electric energy produced by one power plant, or if a trade union confronts one employer. In this case, though one can indicate an upper and a lower limit within which the price will fall, the solution remains indeterminate between such limits.

In all these cases elements intervene which cannot be included in the traditional analysis of market structure, such as the relative strengths of the parties and the politics of the agreements, which are precisely the elements giving rise to determinate solutions in reality. One must also bear in mind that oligopolistic situations are

the most common form of market structure in modern capitalist markets, in which to a great extent industrial development has reduced the number of firms operating in each market to a few.

The grave deficiency of monopolistic competition theory and the indeterminacy of the theory in all cases where oligopoly situations arise, suggest that the revision of the theory of market structure, which occurred during the period between the two world wars, while laying the foundations for a more accurate description of the capitalist market, has not furnished for this market a theory in the proper sense of the word. One must also add that the clash between adherence to reality and theoretical inconsistency, present in the theories that we have examined, raises doubts as to whether a theory of capitalist markets is possible, given the disorder which makes it impossible to extend Sraffa's criticism in any constructive way.

5. *The theory of games*

An alternative line of attack to the unsolved problem of finding an equilibrium solution for the oligopolistic market and for bilateral monopoly was attempted after the Second World War using a new analytical tool—the 'theory of games'—instead of the traditional approach based on cost and demand curves.

Game theory, formulated in 1944 by Von Neumann and Morgenstern on the basis of ideas put forward by the former in 1928, is a rationalization and extension to other conflict situations of the behaviour of players of parlour games.

In a game there are a number of players, each of whom attempts to achieve an objective which is incompatible with those of the other players, but the extent to which each player succeeds in reaching his objective depends not only on his behaviour, but also on the behaviour of all his opponents. This means that, as in fact happens in oligopoly, each player, in deciding which course of action to take, should take into account what is expected of him by his opponents and the way in which his opponents will react. The essential result of the theory of games is that, under certain circumstances, it is possible to find a determinate solution for problems of this type. The theory shows, in fact, that in a 'zero sum game'—that is a game between two players in which the gains of one player are exactly equal to the losses of the other player—there is, for each player, a strategy (more precisely a mixture of strategies over a

sufficiently large number of games) which gives him the mathematical expectation of a gain not less than, or of a loss not greater than, a certain particular value. It also shows that, if the players actually behave in this way, then those expected gains and losses are actually realized and the game has a determinate solution.

The theory of games has brought with it impressive mathematical developments and the work of Von Neumann and Morgenstern is, at least from the mathematical point of view, very interesting. But, from the economic point of view, this theory does not seem to have much relevance because the behaviour of the oligopolists appears to be generally very different from that of players of parlour games, and the similarity decreases even more when the behaviour of the 'players' is defined in very rational terms. To date, there have been no serious attempts to apply the theory of games to actual market problems, or to economic problems in general. But the attention of the economists has been attracted by the interesting mathematical analogies between the theory of games and 'linear programming', another technique of analysis which will be examined in Chapter 8 below.

BIBLIOGRAPHICAL NOTE

P. Sraffa's original article, 'Sulle relazioni fra costo e quantitá prodotta' (*Annali di Economia*, 1925), was republished in English, with some revisions, as 'The laws of returns under competitive conditions' (*Economic Journal*, 1926).

The basic references on the static theory of imperfect competition are J. Robinson, *The economics of imperfect competition* (Macmillan, 1933), containing also a discussion of monopoly, and E. H. Chamberlin, *The theory of monopolistic competition* (Harvard University Press, 1933, 7th edition 1956). On a possible way of introducing the level of advertising expenditure in the analysis, see N. S. Buchanan, 'Advertising expenditures: a suggested treatment' (*Journal of Political Economy*, 1942).

Monopoly and duopoly are treated by A. Cournot in Chapters 5 and 7, respectively, of *Recherches sur les principes mathématiques de la théorie des richesses* (1838), translated by N. T. Bacon as *Researches into the mathematical principles of the theory of wealth* (Macmillan, 1897). F. Y. Edgeworth's theory of monopoly and duopoly is in 'La teoria pura del monopolio' (*Giornale degli economisti*, 1897), republished in English in *Papers relating to*

political economy (3 volumes, Macmillan, 1925).

On the demand conditions facing the oligopolist, see R. L. Hall and C. J. Hitch, 'Price theory and business behaviour' (*Oxford Economic Papers*, 1939), and P. M. Sweezy, 'Demand under conditions of oligopoly' (*Journal of Political Economy*, 1939), who arrived simultaneously at the definition of a 'kinked' demand curve. For a general discussion of oligopoly and bilateral monopoly, see W. Fellner, *Competition among the few* (Knopf, 1949).

The basic reference on the theory of games is J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior* (Princeton University Press, 1944, 3rd edition 1954. On the treatment of the imperfect market as a 'game', see O. Morgenstern, 'Oligopoly, monopolistic competition and the theory of games' (*American Economic Review*, 1948). For an alternative approach, also in terms of struggle between firms but opposed to the game-theoretical treatment, see P. Sylos-Labini, *Oligopolio e progresso technico* (1956, 3rd edition 1961) translated from the Italian by E. Henderson as *Oligopoly and technical progress* (Harvard University Press, 1962, 2nd edition 1969); see also J. S. Bain, *Barriers to new competition* (Harvard University Press, 1956) developed along similar lines, and F. Modigliani's review of both books in 'New developments on the oligopoly front' (*Journal of Political Economy*, 1958).

Chapter 5

KEYNES

1. *Competitive equilibrium and full employment*

Parallel with the revision of the theory of market forms, another criticism was levelled at the theory of economic equilibrium. This criticism concerned the statement that a competitive economy, left to itself, will automatically achieve the full employment of resources in general, and of the labour force in particular. This critique culminated in the *General theory of employment, interest and money* published in 1936 by the English economist J. M. Keynes, one of the major figures in the whole history of economic thought. One could say that, as the theorists of imperfect competition attempted to demonstrate the possibility of a non-competitive equilibrium, so Keynes set out to show the possibility of an equilibrium without full employment. The stimulus to this type of analysis came from the sequence of international events which started with the great crisis at the end of 1929 and involved to a greater or lesser degree all industrialized countries. The circumstances of the crisis naturally cast considerable doubt on the tenet of the traditional theory that an economic system always tends towards full employment. This, in fact, was contradicted by rising mass unemployment everywhere, and especially by the difficulties that were met when trying to eliminate it. These difficulties led one to think that full employment is not necessarily one of the characteristics of an economic system in equilibrium.

There had been other attempts, well before Keynes, to challenge the belief that an economy automatically tends towards full employment. All these attempts were essentially aimed at the formulation of a theory capable of explaining the periodical recurrence of economic crises. The main names here are Malthus, Sismondi and Hobson, not to mention Marx and all the Marxist school. But, leaving Marx aside for the moment, it must be said that, however correct the theses of these economists might be, the analytical tools that they used to prove them were invariably so inadequate that

the orthodox theory of employment was never seriously challenged. We had to wait until Keynes appeared on the scene to find a theoretically adequate critique of the equilibrium theory as far as the problem of full employment was concerned.

2. *The 'classical' theory of employment*

In order to fully appreciate Keynes' theoretical contribution, one has to reconsider the equilibrium theory by looking at it from his particular standpoint. In order to do this, let us distinguish between two statements contained in the traditional or 'classical' theory, as Keynes improperly called it. The first is that it is always possible to find a level of national income equal to the full-employment product. The second is that there are, in the economic system, certain mechanisms which tend to move the economy towards this full employment position.

The first statement is based on the so-called 'Say's Law', from the name of the French economist who first formulated it at the beginning of the nineteenth century. This law states that supply creates its own demand, a contention that can be interpreted in two different ways. One is that Say's Law simply means that the demand derives from the income generated by production and distributed to all those who have contributed to it and, therefore, to the determination of the supply. Interpreted in this way, Say's Law is indisputable. But even in Say's own times, orthodox economists interpreted his law as meaning, more rigidly: whatever the volume of production, the aggregate value of demand is exactly equal to, and therefore cannot be less than, the value of all goods produced, that is, of aggregate supply. It was naturally admitted that there might be occasions when the composition of the aggregate supply is not exactly the same as the composition of the aggregate demand, so that there might be insufficient demand for one particular good, accompanied by an excess of demand for some other good. However, these discrepancies are promptly corrected by movements in relative prices. What was denied was the possibility of an excess of aggregate production, that is, the possibility that, at certain levels of production and income, aggregate demand might be insufficient to absorb all the production. With this more rigid interpretation, Say's Law is much less obvious than it might appear at first sight. We can in fact envisage a situation in which the level of income is very high and total expenditure is less than income; such a level of

income would clearly not be the equilibrium level, and the market prices, left to themselves, would not tend to achieve it. As a matter of fact, Say's Law, although almost universally accepted, was never explicitly proved.

At the time when Keynes tackled the problem, however, the following justification of Say's Law could be constructed on the basis of the contributions of various orthodox economists. With regard to the specific problem of full employment equilibrium, it is convenient to subdivide aggregate demand into two components: the demand for consumer goods and the demand for investment goods. Consumer demand does not pose particular problems; the problems arise in connection with investment demand. The latter derives, in fact, from that fraction of income that, not being consumed, is said to be 'saved'. The problem is to find whether there is some mechanism which insures that the fraction of income saved is always equal to the demand for investment goods; that is, whether there is equality between saving and investment for any level of income. According to the 'classical' economists, this mechanism exists and consists of the movements of the rate of interest. We know that the profitability of an investment project is determined by the market through the comparison of the rate of return on the investment and the going interest rate: it is profitable to invest if the rate of return is not less than the market rate of interest. It is therefore clear that, if the rate of interest falls, more investment projects become profitable and, consequently, a reduction in the market rate of interest is accompanied by an increase in the level of investment. According to this theory, investment is very sensitive to changes in the interest rate, so that there is always a level, perhaps very low, of the rate of interest at which the volume of investment can absorb all the saving, however large this may be. Therefore, even assuming that saving grows when income grows, perhaps even increasing more than in proportion to income, the possibility of generating a volume of investment of the same magnitude guarantees that, for any level of income, there is an equal level of aggregate demand. Consequently, any level of production, including the one corresponding to full employment, is possible in the sense that it can be absorbed by the market.

In addition to all this, the 'classical' theory maintained that full employment income and output is not only possible, but is the equilibrium level of income and output in the sense that the market spontaneously tends to achieve it. In order to prove this point, another mechanism was resorted to. It was stated that, so long as the wage rate is kept equal to the value of the marginal productivity

of labour—that is to the value of the output of the last worker employed—there is always an incentive to employ all the workers available, because the cost of an additional unit of labour is never more than its product and it is therefore profitable to employ it. The problem of unemployment could be then explained only by attritions or by unwarranted interference with the free play of competition. If there are obstacles to the movements of the rate of interest, or the trade unions require a wage rate unmatched by the marginal productivity of labour at full employment, then, and only then, can the phenomenon of unemployment arise.

3. *The Keynesian theory*

We shall illustrate Keynes' critique of the 'classical' theory taking into account not only his own work, but also the clarifications and extensions subsequently made by many other economists. Among the most important of these authors are Hicks, Meade and Harrod in Britain; Klein, Modigliani, Lange and Alvin Hansen in the United States. In presenting the Keynesian theory, it is convenient to start with three propositions.

The first proposition is that saving, while a function of the interest rate (in the sense that the lower the interest rate the greater the flow of saving) is first and foremost a function of income. Keynes also emphasizes how, in modern industrial societies, consumption does indeed grow with income, but less than in proportion to it. As a result, as income approaches the full employment level, the fraction which is not consumed, i.e., saving, becomes larger and larger both in absolute terms and as a fraction of income.

The second proposition, suggested by empirical observation, is that there is a limit below which the market rate of interest cannot fall. To explain this, Keynes resorts to a conception of the interest rate according to which this rate is determined not only by the 'real' phenomena pertaining to the demand and supply of money. This approach contrasts with the 'classical' insistence on the 'real' aspects, of which the most accomplished example is probably Fisher's theory presented above in Chapter 1.

When dealing (in Chapter 1) with the relationship between the theory of general economic equilibrium and the theory of money, we remarked on the tendency, which prevailed before Keynes, to make a net distinction between the real and the monetary aspect of economic phenomena, on the assumption that the demand for

money derives only from the need to carry out transactions. As a result, all the 'real' variables, and the interest rate among them, are independent of the monetary variables. Keynes, on the contrary, considers that, alongside the transactions demand for money, there is a demand for money due to 'speculative' and 'precautionary' motives. This additional demand comes from individuals who wish to hold liquid funds in order to be able to buy earning assets at the precise time when this operation is most profitable, or simply in order to be able to face unforeseen circumstances in business or family life. It is clear that holding liquid funds involves a cost, corresponding to the loss of income that would be forthcoming if such funds had been used to buy earning assets. Therefore, the higher the rate of interest, the higher the opportunity cost of holding money, from which it follows that, as the market rate of interest falls, the demand for money for purposes other than that of making transactions increases. From a monetary point of view, the interest then appears as compensation for the *loss of liquidity*, and the demand for liquidity will be all the greater the smaller the compensation: the equilibrium rate of interest is the rate in correspondence with which the demand for money is equal to the stock of money put in circulation by the monetary authorities.

As pointed out more often by the commentators on Keynes than by Keynes himself, this approach does not imply that the interest rate is a purely monetary phenomenon; what it implies is that the interest rate, like all the other variables of the system, is jointly determined by 'real' (saving and investment) and 'monetary' (demand and supply of money) mechanisms. The theory of the rate of interest provides the link between 'real' and 'monetary' economics, thus putting an end to the traditional separation of the two. It must be said, however, that the process of integrating the real and monetary aspects of economic theory, which could only be done properly by extending the theory of general equilibrium, has not made much progress. No doubt the difficulties met by the theory of general equilibrium in dealing with capital and the rate of interest, as mentioned in Chapter 1, hampered any further development in the monetary direction. Within the context of the Keynesian theory, what matters is that consideration of the monetary factors determining the rate of interest explains why there is a lower limit to the value of this rate. Starting from the fact that the lower the rate of interest the larger the amount of wealth that people wish to keep liquid, Keynes points out that there is generally a value of this rate sufficiently low for people to be prepared to hold any amount of money. So low, that is that there is no incentive for people to

give up keeping all their wealth in liquid form. It is clear that, if the demand and supply of money have to be equal in correspondence with the equilibrium rate of interest, this rate cannot fall below that level at which the demand for money is unlimited.

The third proposition, emphasized mainly by the economists of the Keynesian school (Alvin Hansen, Klein), points to the fact that investment—although negatively related to the rate of interest—has little sensitivity to movements of the interest rate when this rate is already low.

In the light of these three propositions, the equilibrating role attributed by classical economists to the rate of interest can be criticized as follows. As the market rate of interest falls, investment becomes increasingly rigid in relation to the interest rate itself. It could then happen that, in order to have a volume of investment sufficiently large to absorb the volume of saving corresponding to a situation of full employment, the interest rate would have to fall below the minimum level compatible with equilibrium on the money market. Some Keynesian economists have gone even further and maintain that, even if the operation of the money market did not set a lower limit to the interest rate, the volume of investment which is profitable when the rate of interest is equal to nought could still be less than the full employment saving. If this were the case, the level of income corresponding to full employment could never generate an equivalent volume of demand; this level of income would therefore be unfeasible and full employment would not be possible. Naturally the producers could make mistakes. They could, that is, anticipate a certain level of aggregate demand and plan output accordingly; then, if, for the reasons given above, actual demand turned out to be less than expected, the level of production would appear to be too high, sales would not cover costs and a situation of generalized over-production would arise.

Under such conditions, even the classical argument concerning the wage rate cannot stand. Keynes points out that trade unions normally do not accept a money wage rate less than a certain minimum. The 'classics' would say that, if this minimum is higher than the money value of the marginal productivity of labour at full employment, then this is what causes unemployment, which could be eliminated by simply cutting the money wage rate. Keynes showed that if saving and investment cannot be equated at full employment, then a cut in the money wage rate would only result in a fall of the general price level. The real wage rate would, therefore, remain the same or fall by less than the money wage rate and generally not fall enough to restore full employment. In other

words, the entrepreneurs would lose on the demand side whatever they gained on the cost side. If we could suppose that investment can always be increased by the required amount, then Keynes could be criticized as follows. Let us assume, for a moment, that the price level does not fall when the money wage rate is cut, so that the real wage rate falls and employment rises. Aggregate consumption may not increase at all and would certainly increase less than in proportion to total employment, because the larger number of persons employed would now be paid less individually. On the other hand, the additional employment would increase production, and the additional output could be partly absorbed by additional consumer demand, but the rest would have to be absorbed by additional investment demand, which, according to 'classical' theory, can always be stimulated by an appropriate movement in the rate of interest. The initial hypothesis that the price level does not fall would then be legitimate, and the economic system would have then reached a new equilibrium with a higher level of employment. But this mechanism does indeed presuppose that investment can always be increased by the desired amount. If this, as Keynes believes, does not happen, then it is no longer possible to keep the price level constant: a fall in the money wage rate would have deflationary effects, and we would be back where we started.

The conclusion of Keynes' reasoning is that the equilibrium level of income is not necessarily the one corresponding to full employment. What remains to be seen is whether the conditions that make full employment equilibrium impossible really occur in modern economies. This question is legitimate, because, by itself, the Keynesian analysis only shows that under-employment equilibrium is *possible*, leaving open the question as to what are the circumstances in which unemployment would actually occur. The Keynesian analysis is, therefore, a considerable step forward, compared with the classical analysis of capitalism, but still needs to be supplemented by a study of what it is that actually determines the level of investment and, consequently, whether or not aggregate demand will be equal to full employment income.

Keynes' great strength was that he provided a theoretical framework, within which many older ideas and theoretical developments —which before him appeared to be unrelated to one another— could be co-ordinated to form a unitary scheme sufficiently complete to explain what happens in real life. There are two main streams to this extension of Keynes' work. The first is concerned with the theory of the trade cycle, the second with the long-term tendencies of the capitalistic system. We shall examine here only

the first of these developments; the second will be dealt with later, in connection with the theories of economic growth and stagnation.

4. *The trade cycle*

Throughout the whole history of economic thought the trade cycle has been subjected to close scrutiny, but never before Keynes could an adequate explanation of this phenomenon be found, because of the previously mentioned lack of a serious refutation of Say's Law. Limiting ourselves to twentieth-century authors, we must mention the names of the Russian Tugan-Baranowsky, the Frenchman Aftalion, the German Spiethoff, the Englishmen Pigou, Hawtrey and Robertson, the American Mitchell, and of course, Schumpeter, whose ideas we have already examined. Each of these economists could throw light only on particular aspects, though very important ones, of the cycle; what was lacking was a theoretical framework within which the various elements of the jigsaw could be fitted together.

The lack of integration between these individual contributions showed up most clearly in the difficulty these authors encountered in their attempts to explain the so-called 'upper turning-point' of the cycle; that is, the transition from a phase of expansion to a phase of contraction. This transition, because of the sudden and abrupt way in which it occurs, is referred to as the 'crisis'. Before Keynes, these crises were explained by looking for causes *outside* the mechanism of capitalistic expansion. Such causes were identified either in the exhausting of some original factor of production—which in this way set a physical limit to the expansion of production—or, more frequently, in monetary factors—namely the impossibility, under Golden Standard rules, of expanding credit beyond a certain limit determined by the amount of golden reserves available. Because of this attitude, many valuable ideas contained in early studies of the cycle were wasted, for they could not be used to explain the most important aspect of the cycle, the 'crisis'.

As we will see more clearly later, what allows these ideas to be co-ordinated and fully utilized in a Keynesian analysis, is the formulation of the problem in terms of aggregate demand. Two elements have directly contributed to the systematic construction of a theory of the trade cycle. One is the behaviour of consumption in relation to income. We have already seen how, according to Keynes, consumption grows less than in proportion to income, and, there-

fore, as income grows, saving grows not only in absolute terms but also as a fraction of income. This means that, as income grows, the problem of finding a volume of investment sufficiently large to provide an adequate level of aggregate demand becomes more and more difficult. Keynes' other idea—which, incidentally, had some notable precedents in Wicksell and Fisher—concerns the factors determining the decision to invest. We have already said that an investment is profitable if its rate of return is greater than, or at least equal to, the market rate of interest; we must now add that the rate of return depends, in turn, on the future yields expected from the particular type of capital good made available by the investment, and on the cost of the capital good itself. The rate of return on an investment is in fact nothing other than the ratio between its net yield per unit of time and its initial cost. As these yields would be forthcoming over a number of future periods, it is clear that the idea of a rate of return implies long-term expectations.

The question to be asked is the following: how do these expectations vary over time and, consequently, how does the volume of investment—which appears to be profitable in the light of these expectations—vary in relation to saving? To answer this question, modern trade cycle theory mainly makes use of three ideas that can be found in pre-Keynesian analysis. The first concerns the fact, pointed out by Aftalion, that if the stock of capital is to grow faster than the labour force (in order to keep filling the widening gap between income and consumption), then the productive process has to become more and more capital-intensive, and this requires an appropriate rate of technical progress, in the absence of which the opportunities for substituting capital for labour would be much limited. The second idea is Schumpeter's, according to whom the innovations made possible by technical progress tend to *crowd* in certain periods, instead of being uniformly distributed over time. The third idea, finally, concerns the so-called 'acceleration principle', already referred to by Aftalion and later developed more fully by the American Clark. According to this principle, some investments—called 'induced' investments—depend on the rate of change of income, meaning that some of the investments are decided by the entrepreneurs with the purpose of satisfying a demand which is expected to increase at a certain rate. The other investments—those, that is that depend on the opportunities offered by technical progress, rather than on the rate of change of income—are called 'autonomous' investments, to distinguish them from induced investments.

Let us see, now, how these ideas can be grafted on to the

Keynesian construction, with a view to arriving at a theoretically satisfactory explanation of the cycle. A major attempt at this integration has been made by the American economist Alvin Hansen, according to whom the process through which a phase of expansion exhausts itself and ends in a crisis can be described as follows. As income approaches its full employment level, the gap between income and consumption tends to widen, both in absolute and in relative terms. At the same time the incentive to invest becomes weaker and weaker due to two sets of causes. In the first place, as productive resources are used more and more intensively, the rate of growth of income falls from the very high rate previously allowed by the presence of resources not fully utilized and, therefore, available for further production. As a result, that part of total investment which is related, via the acceleration principle, to the rate of increase in income, will also decline. In the second place, accelerated exploitation during the early stages of the expansionary phase of the investment opportunities offered by technical progress, begins to reduce the scope for substituting labour with capital. Another consequence is that autonomous investment becomes less innovating in character and the additional output generated by it will consist more and more of goods that are already known, having been available on the market for a considerable length of time. Therefore profit expectations tend to deteriorate. What is happening is that, just when investment ought to be increasing (in absolute and in relative terms) to absorb an ever-increasing volume of saving, it tends to slow down: induced investment slows down because of the acceleration principle, and autonomous investment because of the exhaustion of the opportunities offered by technical progress. In such a situation, a sizable increase in income would not be accompanied by a correspondingly large increase in aggregate demand; the expansionary process would therefore become weaker and weaker, until income stopped growing altogether. Once a crisis has taken place, a phase of contraction—that is, of progressive reduction in the level of income—sets in. This process is reinforced by the accelerator, now operating in reverse, because, as income starts to fall, 'induced' investment becomes negative and, as a result, there is no more demand for capital goods, not even for replacing those that wear out. But even the contraction phase is bound to end when the point is reached where income is so low that aggregate demand is again sufficient for it not to fall any further. This point is reached for two sets of reasons. On the one hand, the gap between income and consumption tends to close. On the other hand, the fact that no positive investment has taken place for some

time means that there are numerous investment opportunities waiting to be exploited; also, as the phase of contraction draws to a close and the pace at which income is falling becomes slower and slower, the negative effects of the accelerator become attenuated. A lower turning point is thus reached and income starts to grow again.

The Keynesian character of this explanation of the cycles is to be found in the fact that all the various elements of the theory are used to explain how the behaviour of aggregate demand determines the time path of income. In this way, the tradition initiated with Sismondi and Malthus, and then frustrated by the absence of a rigorous criticism of Say's Law, was revived and organized into a coherent theory.

BIBLIOGRAPHICAL NOTE

J. M. Keynes' main works are *A treatise on money* (2 volumes, Macmillan 1930) still presented in 'classical' language, and *The general theory of employment, interest and money* (Macmillan, 1936). The reading of the *General Theory* is facilitated by A. H. Hansen's *A guide to Keynes* (McGraw-Hill, 1953), which follows Keynes' text chapter by chapter, clarifying the obscure points and giving examples. An extensive account of Keynes' life and work is that of R. F. Harrod, *The life of J. M. Keynes* (Scribners, 1955).

The literature on the interpretation of Keynes is enormous. We shall mention only two fundamental articles by J. R. Hicks, 'Mr. Keynes and the "Classics": a suggested interpretation' (*Econometrica*, 1937) and 'A rehabilitation of classical economics' (*Economic Journal*, 1957), where the 'classical' theory of equilibrium is reformulated in terms of aggregate demand and compared with that of Keynes; this shows that the 'classical' full employment equilibrium and the Keynesian under-employment equilibrium are extreme cases of a more general model. The most thorough textbook treatment of Keynesian theory is probably G. Ackley, *Macroeconomic theory* (Collier-Macmillan, 1961). A reappraisal of Keynes, emphasizing the dynamic nature of his theory and the dependence of his conclusions on this aspect of the theory, rather than on the introduction of various rigidities in the static 'classical' model, has been recently attempted mainly by R. Clower and A. Leijonhufvud. For a review, see A. G. Hines, *On the reappraisal of Keynesian economics* (Martin Robertson, 1971).

The three building blocks of Keynesian theory—consumer demand, investment behaviour and portfolio preferences—have initiated three important streams of theoretical and empirical research. The development of consumer theory is mainly due to J. M. Duesenberry, M. Friedman and F. Modigliani; for a good review see M. J. Farrell, 'The new theories of consumption' (*Economic Journal*, 1959). The theory of investment has been developed mainly by A. P. Lerner, T. Haavelmo and D. W. Jorgenson; see A. P. Lerner, 'On some recent developments in capital theory' (*American Economic Review*, 1965) and D. W. Jorgenson, 'The theory of investment behavior (in *Determinants of investment behavior*, edited by R. Farber, Universities National Bureau for Economic Research, 1967; mathematically advanced). On the portfolio preferences of wealth-holders and the demand for money, see J. Tobin, 'Liquidity preference as behavior towards risk' (*Review of Economic Studies*, 1958) and D. Patinkin, *Money, interest and prices* (Row, Peterson & Co., 1956).

The pre-Keynesian contributions to the theory of economic fluctuations are reviewed in A. H. Hansen, *Business cycles and national income* (Allen & Unwin, 1951, 2nd edition 1964); but see also A. Aftalion, 'La realité des surproductions générales' (*Revue d'économie politique*, 1909); J. M. Clark, 'Business acceleration and the law of demand: a technical factor in business cycles' (*Journal of Political Economy*, 1917), and Keynes' essay on Malthus in *Essays in biography* (Macmillan, 1933). On the integration of these contributions in a Keynesian framework see, besides Hansen's book, R. C. O. Matthews' *The trade cycle* (Cambridge University Press, 1959).

Chapter 6

NEW TRENDS IN ECONOMIC POLICY

1. *Government's new role*

The critique between the two world wars of the traditional theories of competition and employment gave rise to considerable changes in the theory and practice of economic policy. The attitude towards economic policy emerging from classical economics (largely supported by the theorists of economic equilibrium) goes under the name of *laissez faire*. This can be summed up in the proposition that the market forces, left to themselves, will generate a situation that no external interference and, in particular, no government intervention, can improve. It was obviously recognized that it was the government's duty to provide the institutional framework within which private economic activity could be carried out in an orderly fashion; but public intervention in the economic sphere was conceived only as being directed towards satisfying collective needs, such as external defence, justice, education, etc. This conception came to be questioned on at least three counts as a result of the theoretical developments illustrated in the last two chapters.

Firstly, it was denied that the market, left to itself, would necessarily give full employment to all available resources, and to labour in particular.

Secondly, it was stated that, as shown by the new theories of the market, part of the available resources employed were not being utilized in the best possible way, because the distribution of these resources among alternative uses was not the optimum which could be obtained through perfect competition only. The allocation of resources brought about by imperfect competition, monopoly and oligopoly is, in fact, often far from the optimum.

Thirdly, the increasing diffusion of monopolistic forms throughout the economic system led one to think that the distribution of income would tend to become more and more uneven. We already know that, according to Robbins the economist cannot pass judgement on the distribution of income. Nevertheless, it is clear that the problem of income distribution has great relevance for eco-

nomic policy; fiscal policies, in particular have always attempted to attenuate the less socially acceptable situations. But now the Keynesian analysis showed that there is a strictly economic side of the distributive problem which had not been realized before. The point is the following. Both empirical observation and *a priori* reasoning show that the propensity to save of individuals with high incomes is higher than the propensity to save of individuals with low incomes, and it is also obvious that all undistributed profits are not consumed. If monopolistic forms become widespread in a certain economy, so that an increasing portion of its national income consists of monopoly rents, the share of income that goes to the richer classes and the share retained within the firms (as undistributed profits) would both increase. As a result, the amount of saving in such an economy would presumably be greater than it would have been under competitive conditions. Applying the Keynesian analysis, one would then have to conclude that full employment would become more difficult to achieve. This consideration shows that the distribution of income is not only a social problem, but also a strictly economic one, because of its effect on the level of employment.

Each of these three considerations gave rise to a change in previously held views of economic policy. The most important of the changes was the one concerning the problem of unemployment, because only with regard to this problem was it possible to formulate an organic and effective economic policy. The results were not equally good as regards the other two problems, that is, the bad utilization of resources and the unsatisfactory distribution of income in a non-competitive situation. The reason for this is quite clear. As we shall see, a full employment policy requires public intervention on a much larger scale than would have been acceptable by the *laissez faire* philosophy but, on the other hand, it does not really modify the structure of the capitalist market economy, its purpose being that of revitalizing and supporting this type of economy rather than changing any of its fundamental aspects. On the contrary, any policy measure that seriously attempted to modify the effects (on allocation of resources or distribution of income) of monopolies would impinge much more substantially on the private ownership of capital itself.

For the time being we shall confine ourselves to an examination of full employment policies. An opportunity to look at anti-monopolistic and redistributive policies will arise later, in connection with the long-term tendencies of capitalism.

2. *The Keynesian fiscal policy*

In accordance with the results of the theory, the aim of Keynesian full employment policies is that of generating a volume of expenditure sufficient to ensure full employment. To understand this we must abandon the hypothesis made so far that total expenditure consists of two elements only, private expenditure on consumer goods and private expenditure on investment goods. We must now add a third element: public expenditure. We shall continue to assume, however, that we are dealing with a closed economy; later we shall see how the Keynesian analysis can be extended to the problems of international trade.

The general principle of Keynesian fiscal policy is the following. Suppose that private expenditure, both for consumption and for investment, is on the whole insufficient to ensure a level of income corresponding to full employment. Suppose also that these two items of total expenditure cannot be increased because, on one hand, consumer expenditure is conditioned by a certain distribution of income, which cannot be changed very easily in the short run, while investment, on the other hand, depends on other factors equally difficult to modify, such as the opportunities offered by technical progress. Under such circumstances the additional expenditure necessary to fill the gap between the full employment level of income and the level of income generated by the automatic operation of the market would have to be public expenditure.

It is evident that this represents a radical change in the traditional views on economic policy. In addition to having to provide certain public services, the government has now in fact assigned to itself the new task of having to guarantee a certain level of national income. If public expenditure is thought of as one of the determinants of the level of national income and employment, the government becomes one of the economic forces of the social system. It is true that the public expenditure manoeuvre does not modify, by itself, the structure of the market, but it does modify its size. It does not affect the mode of operation of private firms, but, by affecting the level of total expenditure, it affects indirectly their scale of operation. These points will become clearer when we have examined some specific aspects of the Keynesian theory of public expenditure.

3. *The level of public expenditure and the multiplier*

The first aspect concerns the problem of determining the level of public expenditure necessary to raise the level of income from the level that it would automatically reach as a result of the operation of the market to the level corresponding to the full employment of the labour force. The problem arises from the fact that any form of expenditure will induce additional expenditures, so that any addition to the level of public expenditure will increase total demand by more than it has increased public demand. This phenomenon, known as the 'multiplier' process, can be easily explained within the Keynesian theory. We have already said that consumption is regarded as a function of income, in the sense that consumption grows when income grows, although less than in proportion to it. Let us assume, for example, that 80 per cent of each increment in income is consumed, and the rest is saved. To simplify matters let us also imagine that we are starting from a situation of underemployment and that total expenditure consists only of private consumption and investment. Suppose, then, that there is a certain amount of public expenditure in addition to private expenditure. As a result, national income will increase by an amount equal to the increase in public expenditure, and this increase in income will be possible because we have assumed that we are starting from a situation of unemployment, so that it is possible to increase production. If nothing else happened, this would contradict the hypothesis that a constant fraction of each increase in income is consumed; the result would be, in fact, that income has increased while consumption has remained the same. To bring the economic system back to equilibrium, consumption must therefore also increase. This means that the initial public expenditure is followed by an increase in consumer expenditure that makes the addition to total demand greater than the mere increase in public demand. The ratio between the increment in aggregate expenditure and the level of public expenditure is equal to the value of the 'multiplier'. This is a purely formal argument, in the sense that it is based purely on the Keynesian model of the economy and not on the observation of real-life economies. But, in so far as this model is a correct representation of reality, this formal argument must be able to explain what happens in real life. What really happens can be described as follows:

Suppose that there is a certain increase in public expenditure. This increase in expenditure will increase income by the same amount, and this additional income will be distributed in the form

of profits, rents, interests, wages and salaries among the various people who have contributed to production. The recipients of this income are, in the first place, the entrepreneurs, the capitalists and the workers of the firms that supplied the goods and services on which public money was spent, and, in the second place, the entrepreneurs, the capitalists and the workers of the whole chain of firms that supplied goods and services to the firms that supplied the government to start with. The second step is that all these people will consume part of their additional income. As happened with the initial public expenditure, this additional consumer expenditure will also generate income for all those who participate in the production of the additional consumer goods and of the productive goods necessary for producing the consumer goods. Part of this additional income will also be consumed and the process will go on and on to infinity. The sum total of these additional expenditures is a well-determined amount. The ratio of the additional expenditure—consisting of the initial expenditure by the government plus the sum of the induced consumer expenditures—to the initial government expenditure itself is in fact equal to the value of the multiplier, which can be shown to be equal to the inverse of the fraction saved of the additional income. If the fraction saved is 20 per cent, that is one-fifth of income, the multiplier is therefore equal to five. For this multiplier process to be possible, it is obviously necessary that, at each stage, there are productive factors available so that the supply of goods can adjust to the additional demand generated by this process. If this were not the case, then the process could not take place in real terms but only in monetary terms; that is income would increase because prices increase, but without any increase in production.

This phenomenon, although an essential part of Keynes' theoretical construction, had already been identified by the English economist R. F. Kahn in an article published in 1931, before Keynes' *General Theory*. It is also worth mentioning that the mechanism of the multiplier, although first presented in relation to a public expenditure aimed at eliminating unemployment, in fact applies to any type of expenditure. We could say that each unit of expenditure has, so to speak, an amount of national income attached to it equal to the value of the multiplier. But, coming back to our initial problem, we can conclude that the level of public expenditure necessary to reach a full employment level of income has to be calculated taking into account the multiplier process. Therefore, the higher the value of the multiplier, the lower the level of expenditure necessary to obtain a certain effect on national income. So, if

the value of the multiplier were five, as in the example above, the additional public expenditure would have to be equal to one-fifth of the increment in income that we want to obtain.

4. *The composition of public expenditure*

The second aspect of public expenditure that has to be considered is its destination. This is one of the points on which the Keynsian approach tends to diverge more sharply from the traditional approach. It does so because the particular goal that is meant to be reached by means of this expenditure is not affected by the destination of the expenditure itself. What really matters, in fact, is that public expenditure creates additional demand. Whether this is done by carrying out public works or by paying subsidies to certain individuals, is of no consequence for this particular problem. Public expenditure could even be completely useless, that is, public money could be spent on things that have no intrinsic value, but it would still have the required effect on the level of national income. The effect required is in fact not that of increasing productive capacity, but, on the contrary, of creating additional demand that will fully utilize a productive capacity already in existence and not fully used. In other words, in so far as unemployment is the result of insufficient demand and not of insufficient productive capacity, what matters is the absolute level of public expenditure and not its content.

Naturally, if there is to be public expenditure, it may be advisable to direct it to things that have some intrinsic utility rather than to something useless. It is always better to spend more public money on, say, public education or certain types of public works that under normal circumstances could not be carried out, than to dig holes in the ground and then fill them up. It is clear, however, that, so long as the problem is that of creating additional public demand to compensate for the insufficiency of private demand, the intrinsic utility of goods and services made available by public expenditure is of secondary importance: the achievement of the main goal does not depend on it.

The third aspect of public expenditure concerns the way in which it is financed, and the answer given to this question within the ambit of the Keynesian theory is another major departure from tradition in the field of public finance. The first point to be kept in mind is that, if public expenditure is aimed at increasing employment, it is effective only if it is really *additional* to the volume of expenditure that would be generated within the economy in the absence of public intervention. It is in fact clear that, if public expenditure were a substitute for other forms of expenditure, then it would not result in an increase of national expenditure and therefore of national income and employment. Now, if public expenditure were financed through taxation, it would become in part a substitute for private expenditure; it would then be not very effective as a means of increasing employment. It is for this reason that Keynesian-inspired fiscal policies rely more on the public debt, a practice known as 'deficit spending'. The public debt is, in fact, a way of employing private savings. Therefore, if public expenditure is financed in this way, it is less likely to become a substitute for private expenditure than if it were financed through taxation.

The public debt, which in traditional public finance was regarded as a temporary instrument to which the government would resort when it was short of cash, becomes now an important instrument of public intervention, capable of modifying the level of aggregate economic activity. The main tenet of classical public finance—the balancing of the budget—now comes to be considered as valid only in a full employment situation, and it is regarded as a mistake in a situation of depression and unemployment, when public expenditure not covered by taxation becomes an instrument of economic recovery. On the monetary side, the danger of inflation, traditionally connected with a government deficit, ceases to exist under conditions of unemployment, because in this case there are factors of production immediately available to increase output and real income, thus preventing an increase in the general price level. It has been pointed out that deficit financing is all the more effective as an expansionary measure if a large part of the public debt is absorbed not by individuals, but by banks, because in the first case the purchase of government bonds would result in a reduction in the stock of money in circulation, while in the second case there could be an expansion in the supply of money. There is also no objection to the Treasury financing public expenditure by directly indebting itself with the central bank.

The practice of deficit spending had already been looked upon as a possible anticyclical measure well before Keynes. But it is only with Keynes that this practice found a rigorous theoretical foundation, and the Keynesian justification is relevant not only for anticyclical policies, but also in connection with long-term problems. We shall come back to this aspect of the question when we talk about the problems of economic stagnation.

6. *Foreign trade policy*

Up to now we have presented Keynesian theories and policies under the assumption that we are dealing with a closed economy. It is now time to consider the international aspects too. The first point to be kept in mind is that, once foreign trade is considered, the total demand for goods and services produced in the country under consideration consists not only of consumption, investment and public expenditure, but also of exports. For the same reason, imports have to be subtracted, because they constitute a demand for goods produced outside the economy under consideration. This means that what has to be added to internal demand is the surplus or deficit of the balance of trade, which thus acquires a role analogous to that of investment. To understand this point, think of national income as the sum of three elements: one consumed, one saved, and one absorbed by taxation. The part saved gives rise to an expenditure if it is absorbed either by investment or by that part of public expenditure which is not covered by taxation, or by the surplus of the balance of trade. It is then clear that, provided that all the other items of total expenditure remain the same, a trade surplus would increase aggregate demand while a trade deficit would decrease it. In the first case national income would tend to increase, while in the second case it would tend to decrease.

The effects on national income of a surplus or deficit in the balance of trade are far-reaching, from the point of view of both the theory and policy of international trade. This is one of the most remarkable examples of the influence that the Keynesian approach has had on contemporary economic thought. The problem is very complex and cannot be examined in full here. We shall limit ourselves to the main points.

The first point concerns the classical theory of the way in which balances of payments are brought into equilibrium. According to John Stuart Mill's *Principles* any deficit or surplus in the balance

of payments of a country is automatically corrected through a mechanism which is essentially based on the movement of prices. If a country has, say, an excess of exports over imports, there will be an inflow of gold in payment for that part of the exports which is not covered by imports. As a consequence of this inflow of gold the general price level will increase in the country under consideration, then its exports will tend to decrease, and its imports will tend to increase. The opposite will happen in a country in which there is an excess of imports over exports. Such a country would have an outflow of gold, which would lower the general price level, thus stimulating exports and discouraging imports. These processes would continue up to the point where the exports and imports are in balance in all countries. The validity of this explanation of the way in which external equilibrium is reached had already been questioned. As pointed out by the American economist Taussig in 1928, doubts arose not because balances of payment do not show a tendency to equilibrium, but because equilibrium is reached much more quickly than could be justified purely on the basis of movements in prices. One suspected, that is, that this tendency to equilibrium was originated from a mechanism much more powerful than the one considered by the classical theory. The nature of that mechanism became clear when the results of the Keynesian analysis became available, after the publication of the *General Theory*, thanks to the contributions of several economists such as Joan Robinson in 1937, Harrod in 1939, Metzler in 1942, Kindleberger and Machlup in 1943.

The Keynesian approach to the problem of external balance differs from the classical approach in the same way as it differs from it in the explanation of how saving and investment are brought into balance. The traditional view was that saving and investment are brought into balance by movements in the rate of interest, while for Keynes this balance is reached via movements in the level of income. Analogously, while the Classics expected exports and imports to be equated via movements in prices, for Keynes this equilibrium is reached via movements in the level of income. The mechanism can be described as follows: if a country experiences an excess of exports over imports this gives rise to an increase in aggregate demand and therefore in income. This increase in income must generate an additional volume of saving which, in equilibrium, must be equal to the surplus of the balance of trade. But if we assume, as seems realistic, that imports are a function of income, then that increase in income will be followed by an increase in imports, which will reduce the trade surplus. Vice versa, if a country experi-

ences an excess of imports over exports, then demand will fall and so will income, and saving will fall by an amount equal to the deficit in the balance of payments. But the fall in income will reduce imports and therefore the deficit will be reduced. Metzler and Nurkse have demonstrated that in order to eliminate all surpluses and deficits the movements in income must give rise to appropriate amounts of induced investments or disinvestments. But what we are interested in at the moment is that there exists a re-equilibrating mechanism based on movements of income.

The practical implications of this way of reaching an international equilibrium are far-reaching. Movements in income are in fact less acceptable, in general, than movements in the level of prices. In particular, if a country finds itself with a substantial trade deficit, the cost that it would have to bear in order to eliminate it might be intolerable. It might happen that the reduction in income, and therefore in employment, necessary to re-equilibrate the balance of payments is such that the objective of achieving a balance would appear out of proportion to the cost that has to be paid for it. In such circumstances the country in question would be tempted to break the mechanism that we have described, by cutting the link between imports and income. The deficit could, in fact, be eliminated by resorting to a protectionist policy that reduced imports. This would involve, to a certain extent, isolation from the international market, and the substitution of at least some of the imports with national produce. It is clear that at times when the international market is very unstable many countries would be induced to take this extreme course of action, because it would not be possible, beyond a certain point, to let the level of income and employment depend on international economic fluctuations. This explains the virtual breakdown of the international market that followed the great crisis of 1929.

What about the other proposition of the classical theory of international trade, concerning the advantages deriving from freedom of trade? Ricardo's theory of comparative costs had shown that free trade, by allowing the various countries to specialize in the production of those goods in which they have a comparative advantage, generates a situation of maximum efficiency for the world as a whole. Each country would derive some advantage from such a situation: by trading freely with the rest of the world it would derive more output from its given resources than it would do in a position of isolation. The Keynesian school does not challenge the validity of this classical proposition, which has been revived, extended and, with some qualifications, confirmed in more recent times by various

authors, such as Haberler, Leontief, Viner and above all Samuelson in his famous 1939 article. What the Keynesian school points out is that there may be circumstances in which the alternative is not between using given resources efficiently or inefficiently, but between using them fully or not. In those circumstances it may well be that, even on the basis of Pareto's criterion, an inefficient but full utilization of resources is preferable to an efficient but incomplete utilization of them. This could be the case of a country for which free trade would mean a reduction in income and employment; which means that the conclusions reached by the classical theory of international trade are correct only in so far as all the countries that take part in international trade have at all times sufficiently high and stable levels of aggregate demand. Otherwise the fluctuations and instability of the international markets would induce some countries to adopt an isolationist policy. As Metzler pointed out in 1949, the more effective Keynesian remedies are in solving the problems of internal stability, the less we need the Keynesian theory to understand the problems of international trade.

BIBLIOGRAPHICAL NOTE

The idea that public spending should be used as a means for increasing the level of employment was put forward by J. M. Keynes in *The means of prosperity* (Macmillan, 1933). At about the same time, the same idea was being put into practice by the Swedish government; see B. Ohlin, *The problem of employment stabilization* (Oxford University Press, 1949). The multiplier argument was first presented by R. F. Kahn in 'The relation of home investment to unemployment' (*Economic Journal*, 1931), and then developed by Keynes in Chapter 10 of *General theory*. The practice of deficit spending was theorized by A. P. Lerner in *The economics of control* (Macmillan, 1944).

On the application of Keynesian analysis to international trade see J. Robinson, *Essays in the theory of employment*, Part III (Macmillan, 1937, 2nd edition 1947); R. F. Harrod, *International economics* (Cambridge University Press, 1933); L. A. Metzler, 'Underemployment equilibrium in international trade' (*Econometrica*, 1950); F. Machlup, *International trade and the national income multiplier* (Blakiston, 1943). The classical view of international trade is reassessed by P. A. Samuelson in 'The gains from international trade' (*Canadian Journal of Economics and Political*

Science, 1939).

A good textbook presentation of the post-Keynesian theory of fiscal policy is A. T. Peacock and G. K. Shaw, *The economic theory of fiscal policy* (Allen & Unwin, 1970).

Chapter 7

THE THEORIES OF ECONOMIC STAGNATION

1. *The classical view of the problem of stagnation*

We have seen how the Keynesian theory of income determination can be used to explain the trade cycle. But the same ideas have also been used to revive one of the oldest problems in economic thought, —the problem of whether the capitalistic economy, as a result of its own expansion, tends to approach a final stationary condition referred to as 'stagnation' or 'maturity'.

In order to understand the theoretical and practical importance of the Keynesian interpretation of this problem better, it will be useful to put the whole problem into historical perspective. The question was first raised by the English classical economists, particularly by Ricardo. Ricardo's belief that a final stationary state would be inevitably reached is the reason for what is referred to as the Ricardian 'pessimism'. Ricardo's argument, later developed by John Stuart Mill, is essentially the following: the wage rate tends to be systematically reduced to bare subsistence level. Once the total wage bill is thus determined, what is left of the national income is used to remunerate the capitalists in the form of profits and the landowners in the form of rents. This residual is what the Classics called 'net income'. The distribution of net income between profits and rents is determined by two principles.

The first principle is that, in the Ricardian system, the rate of profit prevailing in the economy is determined by the profit rate of the agricultural sector. Before Marx it was in fact thought that the value of the stock of capital could be reduced to the total wage bill necessary for keeping alive the labour force. Therefore, if we assume with Ricardo that these subsistence wages consist of wheat alone and that agricultural production consists of wheat only, it follows that agriculture is the only productive activity in which a rate of profit can be actually calculated, because profits and capital are physically homogeneous and therefore it makes sense to divide the one by the other. It is also clear that, in agriculture, the profit

rate can be calculated independently of the price system; and the competitive mechanism, on the other hand, will generate a price system such that the profit rate in the other sectors of the economy is the same as in agriculture. Strictly speaking, this idea that the agricultural profit rate determines the profit rate of the whole economy is one that predates the publication of Ricardo's *Principles of political economy*. In the *Principles*, Ricardo abandons the hypothesis that wheat is the only wage good and that therefore the agricultural profit rate can no longer be determined independently of the general valuation problem. Nevertheless, the idea that the productivity of labour in agriculture plays a crucial role is essential for the Ricardian theory of stagnation, as presented in Chapters 6 and 21 of the *Principles*. For this reason we shall maintain the simplification, and relate the Ricardian theory of stagnation to his pre-*Principles* theory of the profit rate.

The second principle is that agricultural production is subject to decreasing returns. This means that the amount of output, i.e., wheat, produced by one worker decreases as the number of workers employed in agriculture is increased, because the additional workers have to use less and less fertile lands, or to apply more and more capital to the same amount of land.

From these two principles it follows that the marginal land, that is the less fertile land under cultivation, does not pay any rent. The whole product of this land is divided between labour and capital, and the net income consists of profits only. Vice versa, the infra-marginal lands pay a rent equal to the difference between the price of the product (equal to the cost of production, including profits, on the marginal land) and their own cost of production. For this reason this is called 'differential' rent. The net income of the infra-marginal lands therefore consists of profits and rents. The principle on the basis of which net income is divided between profits and rents is straightforward: the profit generated by infra-marginal lands is determined by the profit rate of the marginal land; all the rest is rent. If cultivation is extended, productivity of workers employed on marginal land will decrease because of the hypothesis of decreasing returns; then, as wages already at subsistence level cannot fall, profits and therefore the profit rate must necessarily fall. The profit rate will then fall also on all the other lands, to the advantage of land rent. This process will stop when the profit rate has vanished on the marginal land and, therefore, in agriculture as a whole. But, as the general profit rate is determined by the profit rate in agriculture, in a situation like this the profit rate would vanish in the economy as a whole, and income would consist of

wages and rent only. Then the stimulus to accumulate capital would end and the economy would settle down to a stationary state. Obviously, this tendency could be offset by various factors connected with technical progress, which on the one hand could give rise to increasing returns in agriculture and on the other hand could lead to the progressive substitution of agricultural products with industrial products. Ricardo took into consideration the possibilities offered by technical progress, but did not regard them as being sufficient to avoid ending up in a stationary state. He thought that, at best, technical progress could only delay the advent of the stationary state.

Ricardo's thesis was taken up and enriched by John Stuart Mill, who made the interesting statement that the stationary state, far from being regrettable, as for Smith and Ricardo, is on the contrary desirable because only in a stationary state could the urge to accumulate, characteristic of capitalistic expansion, be superseded by a better balance between the activities aimed at increasing material wealth and those aimed at the peaceful and relaxed enjoyment of the pleasures offered by nature and culture. Mill conceived the stationary state not as the end of technical progress, but as a situation in which the progress inevitably deriving from scientific research is primarily dedicated to reducing working hours and the effort of work.

The problem was proposed again by Marx in dramatic terms that contrasted violently with the rather idyllic position of Mill. Marx was in fact well aware of the fact that a stationary position is incompatible with capitalism and that, once such a system reached a stationary position, its end would inevitably follow and it was therefore necessary to substitute it with an economy of a different type. We should also add that, even if Mill's judgement on the stationary state were acceptable, his idea that such a condition could be reached without any institutional change, and especially without any change in the nature of private property, is totally unacceptable. In the case of Marx, his insistence on the idea that the achievement of a stationary state would be a mortal crisis for the capitalist economy derives from considerations entirely different from those of Ricardo. He rejects the hypothesis of decreasing returns in agriculture on the grounds that technical progress is quite sufficient to overcome the difficulties considered by Ricardo. But he accepts the classical idea that the rate of profit is inevitably destined to fall in the course of capitalistic expansion, and he attempts to explain this tendency with the influence exercised by technical progress on the ratio between the amount of capital des-

tined to acquire means of production and the amount of capital destined to acquire labour services. We cannot go into the details of Marx's thesis here because it would take too long, but it must be said that his 'law', according to which the rate of profit would tend to fall, does not stand. Nevertheless, many of the points on which Marx has thrown light in his discussion of this problem can be used to arrive at a theory of the long-term tendencies of capitalism.

2. *Alvin Hansen and the Keynesian theory of stagnation*

In more recent times the problem of the fate of capitalism has been discussed again using the Keynesian approach based on the concept of aggregate demand. Schumpeter too had put forward important ideas on the subject, based on arguments radically different from those of Keynesian extraction. We shall deal with the Keynesian theory of stagnation first and then with Schumpeter's thoughts on the matter.

From a Keynesian point of view, the question of what will happen to the capitalist economy in the long run can only be answered by answering the question of what determines the level of investment in the long run. Only after answering this question is it in fact possible to say whether the level of investment (ignoring short-term fluctuations) that the capitalistic system considers profitable is sufficient to absorb the flow of saving generated by the economy at full employment. The American economist Alvin Hansen classified the factors that influence the level of investment into three categories: increase in population, introduction of technical innovations and geographical expansion of capitalism.

The increase in population had already been taken into consideration as an important determinant of investment by Keynes in a 1937 article in which he pointed out that, between 1816 and 1913 nearly half the increase in the stock of capital in Britain had to be attributed to this phenomenon. In order to understand how the increase in population could be an incentive to invest, imagine a situation in which there is no technical progress and no chance of geographical expansion. Under such circumstances, if the population were stationary and its needs immutable, a saturation point would soon be reached beyond which there would be no further demand for new capital goods. If the needs and tastes of such a population were not immutable, this would obviously require

changes in productive equipment, but the Keynesian school, like Schumpeter, is fully aware of the fact that in a modern economy changes in needs and tastes are never autonomous, but always induced by changes that have taken place in the world of production.

The second stimulus to capital accumulation is technical progress, which does not require much comment after what we have said about Schumpeter's concept of innovation. The possibility of changing methods of production opened up by technical progress allows given needs to be satisfied at less and less cost, while, at the same time, it gives rise to new needs. This is one of the major reasons why a capitalist economy has to renovate and expand its productive equipment and, therefore, accumulate capital.

Thirdly, the expansion of capitalism to new territories has made available new resources for satisfying the growing and changing needs of a population in expansion. This too has made it necessary to accumulate capital.

The central theme of the Keynesian theory of stagnation, as elaborated by Hansen between 1938 and 1941, is that these three factors of expansion have become progressively weaker in more recent times, while at the same time the industrialized economies have been showing increasing saving potential. If this thesis is correct, it is clear that there is a tendency towards an imbalance between investment opportunities and saving potential, as a result of which the economic system would tend, in the long run, to a situation of under-employment. The system would tend to move along a time-path which is below the one allowed by the increasing supply of labour, and in this sense it would be 'stagnant'. This is a concept of stagnation quite different from the classical one. While for Ricardo the final stage of capitalism would be characterized by a complete absence of growth and accumulation, in the theory that we are considering there would be a tendency to grow more slowly than the increase in the labour force would allow.

We must now examine the reasons for which the stimulus to accumulate has become weaker, but we must avoid making the frequent mistake of thinking that this weakening applies to investment in general: it applies only to the profitability of private investment. The stagnationists' thesis has not met with much opposition as far as their references to population growth and geographical expansion are concerned. It is in fact well known that in all developed countries the rate of increase of the population has been falling. As for geographical expansion, what Hansen was referring to was the fact that the U.S. economy has now completely exploited its own territory. More generally, there is no doubt that capitalistic

expansion into under-developed regions is becoming increasingly difficult. The reasons for this will be examined in Chapter 11. For the time being we can safely assume that capitalism, as we know it today, has already reached its geographical boundaries.

We are left with the other source of accumulation, technical progress, on which the controversy between supporters and opponents of the stagnationist thesis has been more lively, even if largely based on a misunderstanding. The stagnationists have never maintained, as their opponents thought, that there has been a slow-down in technical progress in recent times. What they do maintain is that, however fast technical progress may be, the monopolistic structures that characterize modern economies would tend, more and more, to slow down the rate at which technical advances are exploited and utilized for production.

This thesis can be summarized as follows. An innovation due to technical progress generally requires the installation of new machinery and equipment: this causes the firm a loss equal to the value not yet amortized of the old equipment. If the innovating firm operates in a highly competitive market, the reduction in cost deriving from the innovation can be used by this firm to increase its share of the market at the expense of other firms, and the additional profit will more than compensate for the loss on the old capital. Besides, in a competitive regime every firm would be induced to adopt the innovation in order to avoid being excluded from the market. Vice versa, a firm enjoying a position of total or partial monopoly would not have the same incentive to innovate. Firstly, the monopolistic firm would already control all, or nearly all, the market; therefore the initial capital loss cannot be compensated for by enlarging the market. Secondly, the absence or the scarcity of competitors who could, in their turn, introduce innovations, reduces or eliminates the danger of being excluded from the market. Consequently, the monopolistic firm will tend to postpone the introduction of innovations until the old capital stock is completely amortized, a fact that, although not affecting the rate of technical progress, does slow down the rate at which new techniques are effectively utilized in production.

3. *Implications for economic policy*

The stimulus for elaborating a modern theory of stagnation came from the same sequence of events that had solicited Keynes' critique

of the classical theory of employment. We are referring to the events that started with the crisis of 1929, to the very slow recovery that followed, and to the second crisis of 1938, later interrupted by the war, which took place before the economy had fully recovered from the 1929 crisis. For this reason the economic situation of the years between 1930 and 1939 is referred to as the 'great depression'. This appellation reflected the belief, then widespread, that the economic system had entered an era where incentives to private investment were no longer sufficient to keep economic activity at full employment level.

This thesis had practical implications. What it implied was in fact that the Keynesian policies were not to be regarded as temporary anticyclical measures, but as a permanent feature, necessary for ensuring the full utilization of available resources. The government was thus assigned a role even more crucial than the already important one connected with stabilization policies. It is in fact clear that once public intervention becomes permanent the destination of public expenditure can no longer be regarded as being immaterial, and needs to be carefully specified by means of long-term plans.

The Second World War and the period of inflation which followed pushed the discussion of the problems of stagnation into second place. But, just when theoretical interest in the subject was slackening, the actual policies pursued by the governments of most industrialized countries started to revolve in the direction predicted by the Keynesian stagnationists. Since the end of the war, government expenditure has in fact become a large and stable component of the national expenditure of industrialized countries, to the point that it is now almost unanimously recognized that the economic system could not stand a reduction in this expenditure. In these circumstances it is clear that the destination of government expenditure is a problem of the greatest importance. The decision as to whether this expenditure should go mainly into public works, house building, social security, defence, or into helping under-developed countries, can have far-reaching economic consequences. This point will be taken up again in Chapter 12 of this book.

4. *Schumpeter and the fate of capitalism*

We already know that, for Schumpeter, the capitalistic economy is doomed to reach a final crisis, which will require a change in eco-

nomic organization. In this sense Schumpeter's position is closer to that of the Classics, Marx in particular, than to that of the Keynesian school, because he shares the classical view that the crisis of capitalism cannot be resolved within the ambit of capitalism itself, while the Keynesian position is that the system can be kept alive, with a few modifications, provided that there is continuous and substantial government intervention.

Schumpeter's arguments, however, are not all strictly economic, based as they are on the changes of social structure deriving from the evolution of the capitalistic mode of production. The arguments put forward in his book of 1943, *Capitalism, socialism and democracy*, are mainly based on two arguments, one on 'mechanisation of the entrepreneurial function' and one on modifications of the social and economic environment.

The first argument is that modern statistical techniques and computational devices make it possible to forecast demand with much greater accuracy than was possible in the past. Consequently, the adjustment of productive capacity to market conditions can be made on the basis of more precise calculations, thus reducing the importance that the intuition and courage of the entrepreneur once had. At the same time, the tendency to create very large firms, each controlling a very large proportion of the market, brings with it great risks, deriving from the strict interrelation between the effects of decisions taken at various points of the economic system. This has resulted in production and capital accumulation being controlled more and more by teams of managers and economic analysts rather than by individual entrepreneurs. The very close connection that, in the beginnings of capitalism, existed between the individual entrepreneur and the innovation is therefore broken. The innovatory process is reduced to a routine, the economic process becomes impersonal and automatic, and the entrepreneurial function ceases to pertain to specific individuals. Consequently, the industrial *bourgeoisie* loses its importance, in the same way as it loses its historical role of providing from its midst the entrepreneur, the individual innovator.

Another reason for which the entrepreneurial function would tend to become weaker in mature economies is closer to the Keynesian approach to the problem of stagnation, but Schumpeter's conclusions are very different from those of the Keynesian school. Government intervention in highly developed economies tends either to increase the share of public investment in total investment, or to redistribute income in such a way that consumption increases at the expense of saving, as, for example, in the case of the exten-

sive social security schemes adopted in more advanced countries. We have already seen how this tendency can be justified, in Keynesian terms, as being necessary to maintain aggregate demand at a sufficiently high level to obtain full employment. But Schumpeter's own conclusion is that private investment becomes less and less important for the growth of the economic system, with the consequence that the private entrepreneur also becomes less and less important, as he becomes less and less essential to the accumulation of capital.

If the full deployment of the entrepreneurial function on a private and individualistic basis is, as Schumpeter thought, the central feature of capitalism, the development described above will ultimately lead to a profound and irreversible transformation of the economic system, which would gradually lose its capitalistic connotations and eventually become a planned economy—a situation that Schumpeter considered, if not desirable, certainly possible.

BIBLIOGRAPHICAL NOTE

As P. Sraffa pointed out in his introduction to the *Works and Correspondence of David Ricardo* (10 volumes, Cambridge University Press, 1951), in the *Principles* Ricardo abandons the idea that wages consist of wheat only, so that it is impossible to calculate the agricultural profit rate (and consequently the general profit rate) in physical terms. Nevertheless, this assumption is essential to the Ricardian theory of stagnation as presented in Chapters 6 and 21 of the *Principles*.

On the Keynesian theories of stagnation see J. M. Keynes 'Some economic consequences of a declining population' (*Eugenics Review*, 1937), and A. H. Hansen, 'Economic progress and declining population growth' (*American Economic Review*, 1939). See also E. D. Domar, 'The "burden" of the debt and the national income' (*American Economic Review*, 1944).

On Schumpeter, see the bibliographical note for Chapter 3 of this book.

Chapter 8

THE REVIVAL OF GENERAL EQUILIBRIUM THEORY

1. *The rediscovery of Walras*

The 1930s, marked by the devastating criticism of the theory of competitive equilibrium by Sraffa, Chamberlin and Joan Robinson on the one hand, and by Keynes on the other, also witnessed a renewed interest in the problem of general economic equilibrium, which Léon Walras had posed sixty years earlier.

We have already seen how the theory of general economic equilibrium had been formulated in mathematical terms and why mathematical formulation was necessary in order to describe the complex set of relationships existing between so many economic magnitudes. The use of mathematics was not, however, a novelty; on the contrary, it already had a long tradition in the history of economics, the most significant name in this connection being that of Cournot, whose work dates back to 1838. Compared with these earlier attempts the new theories of general economic equilibrium have the characteristic that mathematics are used not only as a convenient and synthetic expository tool, and as a device for ensuring the logical consistency of the assumptions made, but also for deriving logical conclusions from them. If the theory of general economic equilibrium were not expressed in mathematical terms, it could not even be formulated and the very concept of equilibrium could not be fully understood.

2. *The contributions of Hicks and Wald*

A critical reappraisal of Walras' and Pareto's approach to the theory of equilibrium was made in 1939 by the English economist John Hicks, who contributed a more elegant mathematical formulation as well as an attempt to resolve the problem of its 'stability'. This is the problem of determining the conditions under which the

equilibrium configuration is stable; i.e., such that the system automatically tends to go back to it whenever and for whatever reason it has departed from it. It is interesting that Hicks arrives at the conclusion that the stability conditions consist of those characteristics of consumer preferences and of technology that one would normally expect to find in the real world. A methodological aspect of this line of research, initiated by Hicks and continued by many others such as Oscar Lange and Paul Samuelson, is particularly important. The general equilibrium system is a model that attempts to interpret reality; all the essential features of real-life economies must, therefore, either be built explicitly into the model or at least be consistent with it. The model would otherwise be irrelevant. Because in real-life economies the mechanism of demand and supply gives rise to configurations that are generally stable in the sense explained above, it is important to verify the stability of the theoretical model as a check on its degree of realism.

At the same time the mathematical problems raised by the theory of equilibrium had begun to interest, besides economists, professional mathematicians, whose contributions proved to be fruitful even from a strictly economic point of view. This seems to confirm what we were saying before about the necessity of formulating equilibrium theory mathematically. One of the first important contributions by a mathematician was given in 1936 by the Hungarian Wald. He posed a problem which, at first sight, might seem irrelevant to an economist, but which had far-reaching consequences for economic theory. The problem is that of determining the conditions under which the general equilibrium system has economically meaningful solutions. In other words, supposing that there are no problems of computation, would the quantities and prices that satisfy the equations of the general equilibrium system be economically acceptable, that is, non-negative? Here we can make a point analogous to the one made in connection with the problem of stability: since in the real world there are situations of equilibrium, characterized by prices and quantities that are obviously not negative, then the theoretical model can be correct only if it can be proved that it has these essential properties. The solution of this problem by Wald and by an army of mathematicians and economists who tackled the same problem after him required the use of extremely complex mathematical instruments.

Without going into too many details, we can say something about the changes that have to be made in the Walrasian system in order to ensure that its solutions are economically meaningful. All these changes can be given a precise economic interpretation, but the

only one we shall talk about here consists of substituting some of Walras' equations with inequalities. So the equations between the quantity available and quantity utilized of each productive resource have to be replaced by a relation allowing for a rate of utilization equal to or less than full capacity, with the additional condition that if a resource is not utilized at full capacity, then the equilibrium price of its service is equal to nought. Analogously, the equation between the price and cost of each good must be replaced with a relation allowing for the cost to be greater than or equal to the price, with the additional condition that if the cost is greater than the price that good is not produced.

What these changes mean is that the theory of economic equilibrium is now assigned the additional task of determining which resources are scarce and which goods can be economically produced. But the reasons for which some resources, including labour, may not be fully utilized in a general equilibrium model must not be confused with those determining under-employment of labour in the Keynesian model. In the general equilibrium model the under-employment of resources does not result from insufficiency of aggregate demand, but from the fact that the proportions in which these resources are available are incompatible with the structure of final demand and with the proportions in which they have to be combined in production according to the existing technology. The condition that the price for the services of a resource not fully utilized has to be nought is particularly interesting. This emphasizes the close connection between the concept of price in general equilibrium theory and the concept of social wealth on which this theory is based. Social wealth is in fact defined by Walras as all those things that are at the same time scarce and useful, so that, if something is not scarce, it is not part of social wealth and cannot be evaluated in economic terms.

3. *The Von Neumann model*

In 1932, a mathematician, J. Von Neumann, presented an essay to the University of Princeton which approaches the general equilibrium problem from a position outside the Walrasian tradition. The Von Neumann model differs from the Walrasian one firstly because it does not look at the equilibrium configuration of the economic system at one point in time, but at an equilibrium over time. Secondly, in this model there are no 'original' factors of pro-

duction as there are in Walras, where some factors are either non-producible, like labour and natural resources, or have been produced in a period not included in the analysis. In Von Neumann all the productive factors utilized in one period are none other than the goods produced in the preceding period, which is itself included in the analysis. Production is therefore conceived as a perfectly circular process, in which the same goods appear as inputs and as outputs, without any interference from what may be happening outside the productive process itself. But, as there are no 'original' factors, there can be no 'final' consumption. Remember, in fact, that in Walras consumption was made possible by the fact that the owners of original factors could spend part of the revenue from the sale of the services of these factors on the purchase of consumer goods. In Von Neumann, 'consumption' can only be conceived as a particular type of input needed for the production of a particular output: 'labour'.

Von Neumann, therefore, reduces the economic system to a collection of productive processes transforming goods inherited from the past into others, according to given technological laws. These processes have to satisfy two sets of equilibrium conditions. The first is that, in each unit of time, each productive process has to be operated at such a level that the total input of each good used as a factor of production is less than, or equal to, the total output of the same good in the previous unit of time; with the qualification that, if it is less, the price of that good is nought. Secondly, in order to satisfy the condition of perfect competition, the revenue (that is, the total value of all the goods produced) of each process must be less than, or equal to, its cost (that is, the total value of all the inputs), with the qualification that if the cost is greater than the revenue the process is not operated. Cost and revenue, however, occur at different points in time: therefore, in order to make a comparison possible, future revenues have to be discounted at the market rate of interest or, which is the same thing, the costs have to be augmented by the interests.

These two sets of conditions, by themselves, would allow an infinite number of solutions Von Neumann, therefore, imposes the additional condition that the levels of operation of the various processes have to remain always in the same proportion, so that they must all grow at the same constant proportional rate (which can be positive, negative or zero). Given these restrictions, it can be shown that the model gives a unique solution, consisting of the *relative* levels of operation of the productive processes, the *relative* prices, the rate of growth of the economy and the rate of interest.

It can also be shown that the equilibrium rate of growth (which is the highest allowed by technology) is equal to the equilibrium rate of interest (which is the minimum interest rate in correspondence with which a price system without profits is possible).

The Von Neumann model has a very important place in the history of economic thought. We have in fact seen how the Walrasian approach to the theory of general equilibrium met with insurmountable difficulties when dealing with the problem of capital accumulation. If the output of a certain period includes 'new capital goods' in the Walrasian sense, two hypotheses can be made about these goods. The first is to assume that these goods are productive already in the period in which they are produced—in which case the quantities of capital goods available for current production are no longer given and the model is indeterminate; that is, incapable of defining the equilibrium configuration. The second hypothesis is to assume that new capital goods become productive only in the period following the one in which they are produced—in which case it is impossible to ensure that the rate of return is the same for all the various capital goods, and it is therefore impossible to satisfy simultaneously all the conditions of the Walrasian equilibrium. The Von Neumann model avoids this difficulty by renouncing the idea that there are given quantities of original resources, and treating, in fact, all means of production as intermediate goods. The price that has to be paid for this simplification is, in the first place, that the absolute amounts produced of the various goods cannot be determined and, in the second place, the more serious consequence that consumption can no longer be considered as an economic category distinct from production. Labour is, in fact, an intermediate product like all the others, obtained by means of some particular productive processes; these processes, in turn, have as inputs the consumption of various sorts of other goods. Consumption is, therefore, entirely included and explained within the productive system and the inevitability of this result is confirmed by some attempts, in our opinion unsuccessful, to include autonomous forms of demand in the Von Neumann construction.

The Von Neumann concept of distribution also departs from the traditional one, which originated with Walras. In Walras, the various forms of income are defined as remunerations for the services of the various types of original resources. This conception is obviously incompatible with the Von Neumann model, in which—if the rate of growth of the system is positive—the increase in output from one period to another cannot be imputed to any original factor, and has, therefore, the nature of a *residual*, reminiscent of

the classical-Marxian concept of surplus-value.

Another aspect of the Von Neumann model has been the object of recent research. The restriction that the levels of operation of all the processes have to remain in a given proportion may seem, at first sight, completely arbitrary, imposed only to simplify the mathematical presentation of the model. But Samuelson, Hicks, Morishima and others have shown that this restriction is less arbitrary than it appears. Suppose, in fact, that the productive processes were not growing in the same proportion and that, at a certain future date, we wanted to reach certain target levels of production of the various goods. It has been shown that, in order to reach the target in the shortest possible time, the economic system must abandon the time-path it was originally following, move on to the Von Neumann time-path along which all processes will be expanding in the same proportion at the maximum possible rate, follow this path for most of the time, and only abandon it (i.e., change the proportions in which the various goods are produced) when the target is finally in sight. This is known as the 'turnpike theorem' because of the analogy with the case of a motorist who finds it quicker to speed along a motorway (turnpike) until he is sufficiently close to where he wants to go and then take an ordinary road to his destination, rather than travelling all the way by the ordinary road, which may be shorter but is also slower. Given that this is true whatever the initial situation from which the economy starts and whatever the situation we want to end up with, it follows that the Von Neumann time-path is the most efficient from the long-term point of view. Notice, however, that, by bringing in the argument of long-term efficiency, the Von Neumann model does not become any less arbitrary; what is arbitrary now is no longer the time-path, but the final target, which is taken as given and placed outside the competence of economics.

4. *Input-output and activity analysis*

In this section we shall examine two methods of analysis—'input-output' and 'activity' analysis—which, although applicable to a variety of problems, are particularly helpful in solving a general equilibrium problem.

Input-output analysis originated in the United States between 1930 and 1940 in the work of the economist of Russian origin Wassily Leontief. This method of analysis has the purpose of de-

termining what repercussion a change in the level of production in one sector of the economy will have on all the other sectors. In its final form, the problem posed by Leontief can be described as follows: given a certain level of 'final' demand for certain goods (for private or public consumption, for investment or export), how much has to be produced by each industry and how much has to be imported of each good, if the total demand for each good consists not only of the 'final' demand, but also of the 'intermediate' demand deriving from the fact that each industry needs goods produced by other industries as factors of production?

The Leontief model is based on certain simplifying assumptions which make it possible to find a numerical solution to this problem and, therefore, make the model a practical tool of analysis. The fundamental assumption is that the ratios between the inputs and the output of each industry are constant. These ratios, which are called 'technical coefficients' (analogous to the 'coefficients of fabrication' of Walras), can be calculated from a table showing the flows of payments that have taken place between industries over a certain period of time by dividing the amounts that each industry has purchased from all the others (or imported) by the total amount sold of its own output. Tables of this kind are now available for many countries, including the UK. The model is particularly useful as an instrument of national economic planning because, if the hypotheses on which it is based are considered realistic, it guarantees the internal coherence of the plan in the sense that if the plan did not satisfy the equations of the model, there would be surpluses of certain goods and deficits of others which would need to be corrected.

Investments pose a special problem, because the inter-industrial transactions considered in the model described above refer exclusively to goods purchased 'on current account' during the period under consideration. All purchases 'on capital account'—that is, purchases of goods aimed at maintaining or increasing productive capacity, or maintaining or increasing the inventories of raw materials, semi-finished products, etc.—are excluded. It is these types of purchases that constitute investment. The analytical difficulty of taking into account the inter-industry demand for investment derives from the fact that these purchases do not depend, like current purchases, on the *level* of production of the purchasing sector, but on the *rate of change* of this level. This implies that the analysis has to be re-formulated in dynamic terms, by referring not to the values of the various magnitudes under consideration at one point in time, but to their time-paths.

The first version of the Leontief model avoids this difficulty by considering only the total investment demand for the output of each industry, without attempting to allocate it among the various purchasing industries. This means, however, that the demand for investment cannot be related to production and becomes one of those elements that have to be independently determined (either as forecasts or as planned variables) and there is no guarantee that the changes in the level of production and the levels of investment are reciprocally compatible.

In order to eliminate this shortcoming, in 1953 Leontief presented a new version of his model, known as the 'dynamic Leontief model', which takes into account all inter-industrial transactions, whether on current or capital account. The assumption made about transactions on capital account is analogous to the one made about current transactions, namely that there is a constant ratio between the *levels* of the investment expenditures of each industry and the *change* in its level of production per unit of time: these constant ratios are called 'capital coefficients'. Given the time-paths of final demands (which now do *not* include investment) the model determines the time-paths of the levels of production of all industries. This model is difficult to apply in practice because it is statistically difficult to estimate the capital coefficients, and because of the computational difficulties involved. But, from the theoretical point of view also, the model has been subjected to close scrutiny by many economists, who have specified some of its characteristics and developed certain of its aspects. In particular, one criticism that has been made of the model is that, even if the time-paths of final demand are well specified, the model is still indeterminate; that is, it gives an infinite number of alternative time-paths for the level of production of the various industries and we need some criterion to choose between them.

This last problem can serve as a convenient introduction to the topic of *linear programming*, which is a modern technique for the numerical solution of problems of choice. The adjective 'linear' refers to the fact that the criterion of choice consists in finding the maximum or minimum value of a linear function (in fact a weighted sum) of variables subject to linear constraints; it is this linearity that, on the one hand, facilitates the statistical estimation of the parameters and, on the other hand, makes it possible to calculate the optimum solution. Obviously, this technique can be applied to economic problems only in so far as these can be formulated in the appropriate way; that is, only in so far as economic phenomena can be described by linear functions. In the case of production, for

example, this would require constant returns to scale. There is now a substantial section of economic theory which has been formulated in a way that lends itself to the application of programming techniques and goes under the name of 'activity analysis'. This type of analysis is based on the concept of 'activity', defined as the process of transforming some goods into others according to certain fixed ratios.

Activity analysis and its related programming techniques have had many empirical applications to the problems of individual productive units and to the economy as a whole. Many of the applications of the first type have been in the chemical and petro-chemical industries, in the car industry, in the field of transportation and in solving inventory problems. In all these cases it is generally a matter of finding the level of operation of the various activities open to the productive unit that will maximize revenue or minimize costs. The applications of the second type are closely connected with input-output analysis. Problems of choice can be introduced in the Leontief static model allowing, for example, for alternative methods of producing goods, or by considering alternative sources of supply (for example, internal production or imports). As for the dynamic Leontif model, we have already seen that there is a problem of choice built into it, because the given time-path of final demand is consistent with an infinite number of production paths and therefore with different rates of capital accumulation. One could, for example, start by building a large productive capacity at an early stage in order to have a high level of consumption in the future, or one could aim at a fairly constant level of consumption over time. Similarly, the consumer demands for a certain period could be met out of current production, or by running down inventories accumulated in the past.

Linear programming has also been used as a theoretical tool to prove rigorously some economic propositions concerning the problem of efficiency. Particularly important is the proof that for each optimum plan there is a unique and well-defined price system. The way in which this statement is proved is interesting because it could be defined, in a sense, as operational. If one applies the technique of linear programming to the solution of a problem of maximum efficiency—that is, to the problem of obtaining the maximum output from given inputs or a given output with a minimum of inputs —one finds, in the course of calculation, certain values that can be easily interpreted as prices. These prices emerge during the search for the most efficient solution and remain, so to speak, 'attached' to the optimum solution itself, thus constituting the valuation aspect

of the efficiency problem. This, incidentally, is not a phenomenon peculiar to linear programming problems, as it arises whenever we look for the maximum or minimum value of a function subject to constraints, whatever their mathematical form.

We therefore have an interesting phenomenon of *duality* in problems of choice, which can be well illustrated using one of the most interesting applications of linear programming: the application to the theory of general economic equilibrium. Linear programming can be applied, as we have already said, if there are constant returns to scale. If this is the case, it can be shown in the first place that in a competitive equilibrium, given an arbitrary set of prices for the final products and given the constraint imposed by the limited amount of productive resources, the total value of production is at a maximum. In the second place it can be shown that, given the competitive condition that no productive activity must make a profit, there exists a set of prices for the services of the productive resources available that minimizes the total cost of production. Thirdly, it can be shown that the maximum value of production is equal to the minimum value (cost) of the resources. A competitive equilibrium thus appears as a situation in which, given certain quantities of resources, a certain technology and certain prices of final products, the value of production is maximized and, at the same time, the prices of the services of the productive resources are determined in such a way that the value of production is totally imputed to the resources themselves, without any surplus or deficit.

Using another property of linear programming problems, it can also be shown that a competitive equilibrium is efficient in the Paretian sense; that is, it would not be possible to increase the production of one good without decreasing the production of another. But, of course, there is a different efficient set of quantities produced for each set of prices of the final product: the general equilibrium solution is determined only when the demand conditions—that is, the preferences of the consumers—are brought into the picture. What is important here is that in correspondence with each efficient productive configuration there is a set of prices for the services of productive resources: this phenomenon of 'duality' is of the greatest importance in economic planning, as we will see in Chapter 10.

J. Hicks' contribution to the theory of general equilibrium is in *Value and Capital* (Oxford University Press, 1939, 2nd edition 1946). See also Part II of P. A. Samuelson, *Foundations of economic analysis* (Harvard University Press, 1948). The rest of the literature quoted below is generally mathematically advanced; a more elementary exposition can be found in a seminal book by R. Dorfman, P. A. Samuelson and R. Solow, *Linear programming and activity analysis* (McGraw-Hill, 1958).

The problem of the existence of non-negative solutions is tackled by A. Wald in 'Über einige Gleichungssysteme der mathematischen Ökonomie' (*Zeitschrift für Nationalökonomie*, 1936), translated as 'On some systems of equations in mathematical economics' (*Econometrica*, 1951). See also K. J. Arrow and G. Debreu, 'Existence of an equilibrium for a competitive economy' (*Econometrica*, 1954).

J. Von Neumann presented his model in 'Über ein Ökonomishes Gleichangssystem und eine Verallgemeinerung des Brouwrschen Fixpunktsatzes' (No. 8 of Karl Menger's *Ergebnisse eines Mathematischen Kolloquiums*, 1937), translated as 'A model of general equilibrium' (*Review of Economic Studies*, 1944). See also D. Gale, 'The closed linear model of production' (in *Linear inequalities and related systems*, edited by H. W. Kuhn and A. W. Tucker, Princeton University Press, 1956) and *The theory of linear economic models* (McGraw Hill, 1960); J. O. Kemeny, O. Morgenstern and G. L. Thompson, 'A generalisation of the Von Neumann model of an expanding economy' (*Econometrica*, 1956). For an attempt to include some autonomous form of demand in the model, see M. Morishima, 'Economic expansion and the interest rate in generalised Von Neumann models' (*Econometrica*, 1960) and *Equilibrium, stability and growth* (Oxford University Press, 1964).

The 'turnpike theorem' was given its name by Dorfman, Samuelson and Solow (*op. cit.*, Chapter 12) but others—such as Radner, McKenzie and Morishima—have shown that the particular strategy suggested by the turnpike idea is not always optimal. For a simplified exposition of Radner's version of the turnpike theorem, see F. H. Hahn and R. C. O. Matthews, 'The theory of economic growth: a survey' (*Economic Journal*, 1964).

On the input-output model, see W. W. Leontief, *The structure of the American economy, 1919–1939* (Oxford University Press, 1941, 2nd expanded edition 1951), *Studies in the structure of the American economy* (Oxford University Press, 1953), *Input-output economics* (Oxford University Press, 1966). A good elementary introduc-

tion to input-output analysis, covering both the theory and the applications, is C. Yan, *Introduction to input-output economics* (Holt, Rinehart and Winston, 1969).

The basic reference on linear programming is T. C. Koopmans (editor) *Activity analysis of production and allocation* (J. Wiley & Sons, 1951), containing G. B. Dantzig's seminal paper, 'Maximization of a linear function of variables subject to linear inequalities', and the first statement of the fundamental 'duality theorem' by Gale, Kuhn and Tucker. On the application of linear programming to the input-output model and to the Walras-Wald model of general equilibrium, see Dorfman, Samuelson and Solow, *op. cit.*, and the book by Gale.

Chapter 9

THE THEORY OF ECONOMIC GROWTH

by Alessandro Cigno

1. *'Steady' states and 'stationary' states*

We have seen that classical economists expected the economy to tend inevitably towards a 'stationary state'. This was supposed to come about as a result of the diminishing marginal productivity of labour (the output of a worker employed on marginal land), which would have depressed the real wage rate and the rate of return on 'capital' (the wage fund required to keep the labour force alive) down to the point where there is no further incentive to accumulate. But the picture presented by industrialized countries since the Second World War is very different from the gloomy expectations of the Classics. Having learned the Keynesian lesson on the need to control the level of aggregate demand, governments have succeeded in reducing the dramatic booms and slumps of the past to gentle fluctuations about a rising trend of output and capital. Apart from these short-term fluctuations, the tendency of capital and output has been towards growing at the same constant proportional rate, so that the ratio of capital to output has tended to remain constant. But the rate of growth of capital is itself a ratio between investment —the change per unit of time in the stock of capital—and capital; therefore, if capital is growing steadily, the ratio of investment to capital is constant and, the capital/output ratio being constant, investment is a constant fraction of output as well. The labour force has also tended to grow steadily, but at a slower pace than output; output *per capita* has, therefore, grown at a constant proportional rate equal to the difference between the growth rates of output and labour.

An economy in which labour, capital and output grow steadily, leaving the capital/output ratio constant, is said to be in a 'steady state'. Alternatively, a steady state can be defined as a situation in which the rate of growth of output *per capita* is constant and investment is a constant fraction of output. As we have seen, these two definitions are equivalent, and one implies the other, but the second

one is preferred by some because it avoids mentioning the aggregate stock of capital—a concept as hard to define in theory as it is difficult to measure in practice.

This post-war experience has led many to believe that the long-term tendency of industrialized economies is towards a 'steady' rather than 'stationary' state. A steady state can be regarded as a generalization of the stationary state in that the ratio between capital and output remains constant over time; the stationary state is the special case in which the growth rate is zero and, therefore, the absolute values of capital and output are also constant.

2. *The contributions of Harrod and Domar*

Theoretical studies of the properties of steady states now form a very substantial body of literature, which goes under the heading of 'growth theory'. Some empirical work has been done to verify and apply the theory, but as yet this has lagged far behind theoretical speculation. The approach of modern growth theory to the study of the economy is derived directly from Keynes. Economic operators are lumped into large 'sectors'—households, businesses, government—and all goods produced are aggregated into one composite commodity—the national product—which is partly consumed and partly accumulated as 'capital'. But, while Keynes was concerned with the short-run problem of fully utilizing a given productive capacity, capital and labour being fixed, growth theory is concerned with the conditions under which demand and productive capacity can grow hand in hand in the long run. In this respect growth theory is akin to the theories of Ricardo and Marx; the reasons for the difference in the conclusions that they reach must be found in the different assumptions made about the functioning of the economic system.

The classical prediction of a long-term tendency towards a 'stationary state' rested on the twin assumptions of fixed land and given technology. As production is increased, the marginal productivity of labour—applied to a fixed amount of land according to a given technology—must necessarily fall and so does the rate of return on the accumulating wage fund (capital), until this rate is so low that there is no further incentive to accumulate and the economy grinds to a halt. The assumption that land is a scarce factor of production was necessary to justify decreasing returns because the labour force was supposed to vary directly with the wage rate. But,

in a modern industrialized economy, farming is too small an activity for land to be regarded as significantly restricting total production, and the 'welfare state' has cut the link between wages and size of the population; in modern growth theory land is therefore ignored and the labour force is assumed to be growing at a constant proportional rate, determined by demographical and cultural factors only. This change of assumptions has two important implications. The first is that capital can no longer be seen as a wage fund, equal in equilibrium to the number of workers required times the subsistence wage rate; it must be looked upon as the stock of composite commodity required as a factor for further production. The second implication is that—as production is not restricted by available land and the supply of labour does not depend on the wage rate—the rate of return on capital will only fall if the stock of capital grows faster than the labour force, thus making capital relatively abundant; but if the two factors stay in the same proportion the rate of return on capital will remain constant, and would actually rise if labour grew faster than capital. Therefore, there is no inherent reason, even in the absence of technical progress, why accumulation should eventually stop.

The earliest contributions to growth theory are those of the Englishman Roy Harrod and of the American Evsey Domar, who arrived independently at the same conclusions although starting from different standpoints. Harrod's ideas were first presented in an article published in 1939—the same year as Hicks was re-proposing the problem of general static equilibrium—and are particularly interesting because they contain the essential ingredients of most subsequent theories. His own presentation, however, is rather informal and lends itself to different interpretations; the interpretation given below is the most straightforward and the one that we consider the most acceptable on purely logical grounds.

Stripped down to the bare essentials. Harrod's view of the economic process reduces to two hypotheses. The first is that capital and labour have to combine in a fixed proportion dictated by current technology, so that a fixed amount of capital and a fixed amount of labour are required to produce one unit of national product. The second hypothesis is that the community wishes to consume a constant proportion of its income and 'save' the rest. But, as the accumulated stock of the composite commodity is the only available asset, and no explicit distinction between households and businesses is made, the only way of 'saving' in such an economy is by 'investing', i.e. adding to the stock of capital; desired investment too is therefore a constant fraction of the national

product. The rate of growth of the capital stock is then the ratio of two constants: the flow of saving and investment per unit of output desired by the community, and the stock of capital per unit of output dictated by technology. This is also the rate at which output must grow if the stock of capital is to be fully utilized at all times; Harrod calls it the 'warranted rate of growth' of the economy, i.e., the growth rate warranted by the community's readiness to forsake present for future consumption. But as a fixed amount of labour, as well as capital, is needed to produce a unit of output, the rate of growth of output must also be the same as that of the labour force —what Harrod calls the 'natural rate of growth' of the economy— if full employment is to be maintained. Full employment of labour and capital is, therefore, possible only if the 'natural' and the 'warranted' rate of growth are the same, i.e., if the labour force happens to be growing at a rate equal to the desired saving/income ratio divided by the fixed capital/output ratio.

Domar's approach differs from Harrod's in that the stock of capital and the technological laws regulating production are not explicitly mentioned; the focus is, instead, on the dual role of investment. Keynes, in his preoccupation with the short run, had seen investment only as a component of total demand, which, as such, only needed to fill the gap between a given productive capacity and consumption. But, in the longer run, investment increases productive capacity: Domar's problem is finding the conditions under which the effect of investment on productive capacity is exactly equal to its effect on aggregate demand. Like Harrod, Domar assumes that the community wishes to save a constant proportion of its income, so that the *change* in aggregate demand between two periods is equal to the *change* in investment times a constant multiplier—the inverse of the desired saving/income ratio. He also assumes a constant 'productivity of investment', so that the *change* in productive capacity between two periods is equal to the *level* of investment during the initial period multiplied by a fixed coefficient. Obviously, the change in demand will equal the change in productive capacity only if the change in investment divided by the desired saving/income ratio equals the level of investment in the initial period multiplied by its productivity; therefore, the change in investment between two periods must be equal to the initial level of investment times a growth rate, calculated by multiplying the desired saving/income ratio by the productivity of investment. If investment increased at a lower rate (or remained constant), the increase in demand from one period to the next, generated by the multiplier mechanism, would be smaller than the additional productive capa-

city created by investment during the first period. Conversely, if investment increased at a faster rate than the community's propensity to save and the productivity of investment allow, the multiplier effect on demand would be greater than the increase in productive capacity. We thus have the paradox that, if an economy invests too little, it will have excess capacity *as if* it invested too much, and things will get worse and worse if entrepreneurs attempt to eliminate the excess by slowing down investment. On the other hand, an economy that invests too much will have insufficient productive capacity *as if* it did not invest enough, and any attempt to bring capacity in step with demand by accelerating investment will only make demand run further ahead.

Domar's analysis, though more limited in scope than that of Harrod, throws light on a fundamental aspect of the growth process: investment has to increase at a particular growth rate, no more and no less. It can be easily shown that this is none other than Harrod's 'warranted' rate, which in turn must be equal to the 'natural' growth rate of the economy. Domar does not specify how labour and capital are combined in production, but his assumption of constant productivity of investment—that is, of a constant ratio between change in productive capacity and change in capital stock —can have only two alternative implications. One is that, as for Harrod, labour and capital must be combined in fixed proportions; in which case the ratio between capital and output (and, consequently, between their changes) is constant. Alternatively, it could be assumed that labour and capital can be combined in varying proportions, but in that case the marginal productivity of capital (and, consequently, the ratio between change in productive capacity and investment) would fall whenever more capital is used relative to labour and vice versa; labour and capital would then have to grow at the same rate in order to justify a constant productivity of investment. In either case capital, output and labour would have to grow at the same constant proportional rate, and the capital/output ratio would be constant and equal to the inverse of the productivity of investment. Domar's required growth rate of investment is therefore the same as Harrod's 'warranted' rate, and it must be equal to the 'natural' rate for a steady state to be possible.

The most obvious defect in this simple construction is that it will only allow for a zero rate of growth in output *per capita*, because output and labour have to grow at the same rate: the most desirable feature of economic growth is, therefore, inconsistent with the theory. To get round the problem, Harrod defines a 'neutral' form of technical progress that allows output to grow faster than labour without

altering the amount of capital required per unit of output. This form of technical progress—the only one consistent with steady growth—amounts to an all-round improvement in the efficiency of labour and can be easily introduced in the foregoing analysis by the simple device of measuring the labour force at each point in time by counting not actual heads but 'efficiency units', that is, the equivalent number of workers that would haev been required to produce the same amount of output at an arbitrarily chosen date of the past. The 'natural' rate of growth then becomes the sum of two growth rates: that of the labour force measured in natural units, and that of efficiency per worker, or 'rate of technical progress'. The increasing efficiency of labour therefore allows output *per capita* to grow at a rate equal to the difference between the growth rate of output and labour, that is at the rate of Harrod-neutral technical progress.

Two important conclusions emerge from the Harrod-Domar line of thought. One is that the long-term rate of growth of the economy cannot be a target for economic policy; not, at least, for the ordinary monetary and fiscal policies of Keynesian extraction. The only sustainable rate of growth is, in fact, the 'natural' rate of growth, which the government can only hope to influence if it has ways and means of stimulating technical progress. The other is that a situation of sutained growth with full employment of labour and capital is a mythical state of affairs—a 'golden age', as Joan Robinson called it—which can only come about if the rate of population growth plus the rate of technical progress happens to equal the desired saving/income ratio divided by the capital/output ratio, and, of course, if there was full employment to start with. Besides, the 'golden age' path is a very slippery road—a 'knife edge'—and any accidental departure from it would take us further and further away from it: a slow-down in investment would, in fact, cause aggregate demand to lag behind labour and capital, giving rise to increasing unemployment of resources, while an unwarranted acceleration in investment would cause demand to run ahead of productive resources, giving rise to inflationary pressures.

3. *The neo-classical revival*

The inherent instability of the economic system as seen by Harrod and Domar is due to the fact that these authors take the typical features of a steady state—constant ratio of capital to output, con-

stant fraction of output saved and invested, constant rate of growth in labour productivity—as given data, instead of trying to explain them as the results of the operation of the system. Their analysis, therefore, reduces to finding the conditions under which the numerical values of these constants and the rate of growth of the labour force are reciprocally consistent. Much of the research that followed the stating of the Harrod-Domar conditions is an attempt to explain one or other of these data as economic variables.

As Robert Solow points out in a fundamental article published in 1956, 'a remarkable characteristic of the Harrod-Domar model is that it consistently studies long-run problems with the usual short-run tools. One usually thinks of the long run as the domain of the neo-classical analysis, the land of the margin. Instead Harrod and Domar talk of the long run in terms of . . . "the" capital coefficient.' He then sets out to construct a model of an economy which differs from Harrod's only in that the range of techniques available allow capital and labour to be smoothly and continuously substituted for each other in producing the composite commodity. In such an economy the marginal productivity of a factor will rise or fall depending on whether the input of the other factors is increased or decreased, but will stay the same if the factor mix is kept stable. As land is not a scarce factor, this implies, among other things, that if the inputs of labour and capital are increased in the same proportion output will increase in the same proportion too; we therefore say that there are 'constant returns to scale'. Furthermore, competition tends to equate, at least in the long run, the rate of remuneration of each factor to its marginal productivity. The implications are far-reaching.

Firstly, all the important results of economic activity now depend on the choice of factor mix. The larger the stock of capital per unit of labour, the greater the flow of output per unit of labour and the capital/output ratio. Furthermore, the larger the stock of capital per unit of labour, the higher the wage rate (equal to the marginal productivity of labour) and the lower the rate of return on capital (equal to the marginal productivity of capital).

Secondly, satisfying the Harrod-Domar conditions is not a hit-or-miss affair, because the capital/output ratio is variable and can be selected so as to equate the warranted rate of growth to the natural rate of growth. The capital/output ratio, that is, must be set equal to the desired saving/income ratio divided by the natural rate of growth. In a steady state the natural rate of growth and the fraction of income saved are both constant; consequently the capital/output ratio must be constant too. But the capital/output

ratio can only be constant if the stock of capital per unit of labour does not change; therefore, in a steady state, output and wages per unit of labour and the rate of return on capital will be constant too. If there is technical progress, however, the efficiency of labour is continually increasing; then it is only the output and wage rate per efficiency unit of labour that remain constant, while the output and wage rate *per capita* grow at a constant proportional rate equal to the rate of Harrod-neutral technical progress.

Thirdly, the capital/output ratio must vary directly with the desired saving/income ratio and inversely with the rate of growth of the labour force. But the required capital/output ratio determines the proportions in which labour and capital must be combined; therefore, the higher the propensity to save of the community and the slower the growth of the labour force, the greater the stock of capital and level of output per unit of labour, the higher the wage rate and the lower the rate of return on capital. From this we can see that, in the long run, the rate of growth in output *per capita* does not depend on how much the community is prepared to save, but the absolute level of output and income *per capita* do. This has some important policy implications that will be examined later.

Finally, it can be shown that, under some more specific assumptions about the properties of the technology available, a 'golden age' path not only exists, but is also 'unique' and 'stable'. These technological properties in fact make sure that only one particular factor mix is consistent with each saving/income ratio, so that, once this ratio is chosen, the steady-state behaviour of the economy is unequivocally determined. Furthermore, any mistake in the choice of factor mix will be self-correcting. Suppose, for example, that the stock of capital per unit of labour were 'too small', in which case the warranted rate (saving/income ratio divided by capital/output ratio) would be greater than the natural rate. This is the same as saying that output per labour unit times the saving/income ratio is greater than capital per labour unit times the growth rate of labour, which means that the flow of saving and investment per unit of labour employed is more than is required to equip additional units of labour with the same stock of capital as those already employed. The stock of capital per unit of labour would therefore increase, but at a decreasing pace because the productivity of investment would fall as labour became scarce relative to capital. The process would eventually stop when the mix had become 'right', which is when the capital requirements of new workers absorb investment completely and the warranted rate of growth is equal to

the natural one. The same process would operate, in reverse, if the stock of capital per unit of labour were 'too large'. The golden age path does not lie on the sharp edge of a knife, but at the bottom of a trough to which the economy automatically tends to return.

This model projects the image of a well-ordered and stable society, whose prosperity depends on the thriftiness, inventiveness and technical skill of its members, and where everyone is rewarded according to his contribution to the common good: we are definitely in neo-classical territory. The working of such an economy can be described as follows: at the beginning of each period there is a given labour force and a given stock of capital (buildings, machinery, inventories of finished products and work in progress) resulting from past accumulation. In the course of the period competition will ensure that labour and capital are combined in the most efficient way allowed by current technology, and that the product is distributed to the factors as wages and profits according to their marginal productivity. The recipients of these incomes decide how much they want to save, thus determining how much of the national product is consumed immediately and how much accumulated as capital. In the next period the larger capital stock and labour force, combined according to an improved technology, will produce a larger output, part of which will go to increase the capital stock of the period after, and so on.

In a steady state, the percentage increase in output *per capita* from one period to the next is due entirely to technical progress, because the percentage increase in capital is exactly equal to that of the effective labour force and the capital stock per efficiency unit is consequently constant. One might therefore expect the capitalist-entrepreneur to reap the benefits of growth, but what happens in the neo-classical model is exactly the opposite: it is the wage rate that increases with technical progress, while the rate of return on capital stays the same. This can be justified within the Schumpeterian scheme: the aim of the innovator is, indeed, to realize a surplus over the current cost of labour and capital, and this is the mainspring of technical progress, but in the long run the competition between firms bidding for an inelastic labour supply will convert the surplus into higher wages.

It is the long-run perspective again, that allows neo-classical growth theory to treat investment as an accommodating variable, obediently adjusting to the desired saving. In order to distinguish between decision to save and decision to invest, let us bring into the analysis some form of paper asset—a debt instrument issued by the business sector or by some financial intermediary, which savers

can accumulate as an asset instead of directly acquiring real capital. Under perfect competition the supply of this asset will depend on its interest rate, because businesses maximize profits by borrowing up to the point where the marginal productivity of capital equals the market rate of interest. The demand, on the other hand, will grow at the same rate as income because the saving/income ratio is kept constant. Demand and supply are therefore equal in a steady state, where capital and output grow at the same speed. If investment got out of step with saving, there would be excess supply or excess demand for the asset, the interest rate would move up or down, dampening or stimulating investment as the case may be, and the 'error' would be corrected. Similarly, if the community decided to save more or save less, investment would follow suit, the factor mix would gradually become more or less capital-intensive, and the economy would move towards a new golden age appropriate to the new saving/income ratio. The consumer is therefore master of this neo-classical universe: his choice between immediate and future consumption sets the course of the economy, while investment decisions play a passive role.

The problem with the neo-classical explanation of steady states is that it relies so crucially on special assumptions about technology. In the literature about neo-classical growth models these assumptions are summed up in a mathematical equation, called the 'aggregate production function', which relates aggregate output to the aggregate capital and labour employed. The very concept of an aggregate production function has attracted much criticism. Joan Robinson led the attack, as early as 1953, along the following lines: the existence of an aggregate production function presupposes the existence of an aggregate stock of capital, but capital is a collection of heterogeneous goods that can be added together only in value terms. To construct the aggregate capital we therefore need the prices of individual capital goods, which in turn can be derived from the prices of the goods produced only if the wage and interest rates are known. We are then entangled in a circular argument: the aggregate production function cannot be defined without the concept of aggregate capital, for which we need the wage and interest rates; but these rates are assumed equal to the marginal productivities of labour and capital; therefore we need an aggregate production function defining the technical relations between factors and product before we can talk of aggregate capital.

This question of the existence of aggregate capital and aggregate production functions has been taken up more recently by Franklin Fisher and several others, who have looked at it from a different

angle. The way in which they pose the problem is the following: given that each individual firm has a particular 'production function', which determines the maximum amount of output that the firm can produce by using its factors in the most efficient way, can a similar relationship be found for the economy as a whole? The answer would always be 'yes' if output, labour and capital were made up of physically homogeneous elements, and labour and capital could be moved from one firm to another so as to obtain the maximum output for the economy as a whole. But neither of these conditions applies in practice and it turns out that an aggregate production function exists only in some narrowly defined circumstances.

A special case that has been extensively explored in the literature is the one in which output is homogeneous, labour is also homogeneous and perfectly mobile, and capital only is heterogeneous and immobile—but the heterogeneity is due entirely to technical progress. The only form of technical progress that we have considered so far is the one that increases the efficiency of labour, irrespective of whether there is investment or not. This form of technical progress can only derive from improvements in management and organization, and in education if the corresponding expenditure is not counted as investment. What it cannot account for is technical advances tied up with investment, such as the introduction of improved machines and equipment. If this other form of technical progress is accounted for, capital goods of different 'vintages' will have different degrees of efficiency because they embody the state of technical knowledge of the day when they were installed; nevertheless, a capital aggregate can be constructed and the aggregate production function exists because we can measure capital in 'efficiency units' as we did with labour. Unfortunately this turns out to be the only source of capital diversity that can be accommodated within an aggregate model. For this reason, one of the directions in which neo-classical growth theory is moving is towards the construction of partially disaggregated models with heterogeneous capital goods.

4. *The new Cambridge school*

In addition to the logical difficulties met in building a model around an aggregate production function, the neo-classical theory of growth has been charged with being unrealistic because of the

vision of a consumer-dominated society on which it is based. Most of the critics—Kaldor, Pasinetti, Joan Robinson, to mention but a few—are Cambridge economists and have enough in common to justify their being called a school: the 'new' Cambridge school, as distinct from the old one of Marshall and disciples.

We have seen that the fundamental decisions of a neo-classical economy are taken by the consumers and transmitted to the producers via the mechanism of perfect competition. Even technical progress appears to be quite unrelated with business motivations: either it falls 'like manna from heaven' at an immutable rate, or it is embodied in new capital goods—in which case the diffusion of innovations depends ultimately on decisions taken in the household sector, because the flow of investment is determined by desired saving. This picture seems far removed from the reality of modern capitalistic economies, where competition is severely restricted, consumers are little more than pawns of the game and all the important decisions appear to be taken in the boardrooms of large business firms. A simple model, first presented by Nicholas Kaldor in 1957, can serve as an illustration of the alternative growth mechanism envisaged by the new Cambridge school, as it neatly turns upside-down the order of causation of the neo-classical model.

The crucial assumption underlying Kaldor's model is that competition is not perfect and factors are *not* remunerated at their marginal productivity value. The desired stock of capital and the decision to invest are still based on the rate of return on capital, but this rate is not determined by the marginal productivity of capital: as Kaldor points out in a later article, this is '... an *assumed* rate of profit, based on a mixture of convention and belief, which enables entrepreneurs to decide whether any particular project passes the test of adequate profitability'.

The second important point is that technical progress is not automatic: every community has an innovatory potential determined by its native inventiveness and degree of technical sophistication, but the actual rate of technical progress depends on how promptly the business section of the community takes up the opportunities for innovating. Innovating means changing technique of production —that is, in an aggregate context, changing the proportion in which labour and capital are combined. The rate of growth in output *per capita* (or rate of technical progress) therefore depends on the rate of growth in capital *per capita*, and it is impossible to make a distinction between a switch to a technique belonging to the store of current technical knowledge and the adoption of a newly discovered technique: either move is a departure from current practice and, as

such, an innovation. Kaldor postulates that the growth rate of output *per capita* increases smoothly with the growth rate of capital *per capita*, though at a diminishing rate, and calls this relationship a 'technical progress function'. The rationale behind this assumption is the following: imagine to begin with a situation in which investment is just enough to keep up with the increase in the labour force, and to start investing more. Labour productivity will start to rise sharply causing output *per capita* to grow faster than capital *per capita*—technical progress is then 'capital saving'. But, as investment is accelerated, it will tend to exhaust its innovatory potential, the rise in labour productivity will slow down and the rate of growth in capital *per capita* will eventually catch up with that of output *per capita*—at that point technical progress is 'neutral' in the sense of Harrod, because it leaves the ratio of capital to output unchanged. Beyond that point capital *per capita* will grow faster than output *per capita*, and technical progress will be 'capital using'.

The technical progress function thus allows only one common rate of growth of capital and output *per capita*. This is, by definition, the steady-state rate of technical progress, which added to the given rate of increase in the labour force makes the 'natural' growth rate of the economy. But investment is motivated by profit, therefore a particular profit rate is required if capital is to grow at the 'natural' rate; furthermore, desired saving must somehow equal investment. As is typical of Cambridge models, what does the trick is the distribution of income. Profit earners are assumed to have a greater propensity to save than wage earners, so that total saving depends on how income is distributed between capital and labour: the larger the share of profits in income, the greater the fraction of income saved. It can be shown that, within certain limits, it is possible to find a distribution of income such that planned saving and investment are equal, and that the desired ratio of investment to capital —i.e. the warranted rate of growth—equals the natural rate of growth.

The argument runs as follows. Suppose, for simplicity, that saving is a given fraction of profits while wages are entirely consumed, and that the decision to invest is governed by the desire to keep the capital/output ratio in a certain relationship to the rate of return on capital. In a golden age capital grows at the natural rate, thus investment is equal to the capital stock times the natural rate of growth. But investment must also be equal to saving; therefore capital times the natural rate of growth must equal profits times the desired saving/profits ratio, which is the same as saying

that profits per unit of capital must equal the natural rate of growth divided by the desired saving/profits ratio. The rate of return on capital is thus constant and is not determined by the marginal productivity of capital, which does not come into the picture at all, but by the need to have a rate of saving that will allow capital to grow at the natural rate. The rate of return on capital then determines the long-term capital/output ratio, which is obviously constant.

Given the initial capital endowment of the economy and the natural rate of growth, we can now find the path followed by capital in a golden age; from this, given the constant capital/output ratio and the rate of return on capital, we can then derive the paths of output, profits and, as a residual, wages. As in the neo-classical model, we also find that the growth rate of output *per capita* is determined by technical progress only, while its absolute level is directly related to the propensity to save, but the way in which these results are arrived at is very different. According to the neo-classical logic we would start by observing that the fraction of income saved in a steady state is constant, then we would assume that investment adjusts to saving and in so doing determines the appropriate factor mix; this would in turn determine the constant capital/output ratio and output *per capita* on the one hand, and the rate of return on capital and distribution of income on the other. With Kaldor, vice versa, we start off by noting that the steady-state ratio of capital to output is constant and that a particular rate of accumulation is therefore required; then we assume that saving can be adjusted to the required investment by the way in which income is distributed and the rest follows. For Kaldor, therefore, the steady-state distribution of income determines the rate of return on capital and the capital/output ratio, not vice versa as in neo-classical theory. Furthermore, the prime mover in Kaldor's vision of the growth process is not the consumer but the producer, whose '. . . readiness to absorb technical change combined with the willingness to invest capital in business ventures' determines both the rate of growth and the distribution of output; this makes the model Keynesian in spirit.

In the Keynesian spirit is also Kaldor's explanation of the stability of the golden age, but the argument is not as fully developed as in neo-classical theory. According to the latter, a departure from a steady state could be caused by an 'error' in planning investment, which would be brought back into line with saving by a change in the rate of interest. In contrast, Kaldor's line is that saving could fall out of step with investment because of a 'wrong'

distribution of income. But, if desired saving were smaller or greater than planned investment in a situation of full employment, aggregate demand would either exceed or fall short of productive capacity; the price level would then have to move up or down to close the gap, the real wage rate and the distribution of income would change until saving equalled investment, and the economy would be back in the golden age. This reasoning is also consistent with Keynes' explanation of the possibility of a short-run equilibrium without full employment. Prices and the rate of profit cannot in fact change very quickly, therefore the equilibrating mechanism envisaged by Kaldor cannot apply in the short run, when the adjustment of saving to investment will be achieved through changes in production and employment brought about by the multiplier process.

A more sophisticated version of this model was presented by Kaldor in collaboration with Mirrlees in 1961. The fundamental difference between the two versions is that, in the later one, technical progress is assumed to be mainly 'embodied' in new machines and equipment; therefore capital goods of different vintage are not homogeneous and the aggregate stock of capital becomes a dubious concept. The implications are twofold. Firstly, the technical progress function has to be redefined as a relationship between output per worker *operating on new equipment* and investment per worker: as in the non-vintage version this relationship determines the steady-state rate of technical progress. Secondly, the decision to invest cannot be based on a desire to achieve a particular capital/output ratio, but on whether a particular project is sufficiently profitable and not too risky.

A simple rule of thumb used by businessmen to assess whether an investment is profitable or not is to see whether the expected rate of return is at least equal to what, on the basis of experience, is generally regarded as the norm. But it turns out that in a steady state, where everything changes in a predictable way, expectations are always fulfilled, so that the actual and normal rate of return on capital are the same. Since only one rate of increase in investment is consistent with steady growth, the rate of return on capital and the division of income between profits and wages are then unequivocally determined—again without any reference to the relative scarcity and marginal productivity of capital and labour. The other conclusions derived from this model are also substantially similar to those of the earlier version, but it is interesting to find that this way of explaining the process of growth stands up to the test of more realistic assumptions about the functioning of the economic

system.

Much has been made of the fact that neo-classical growth theory is built around the concept of a production function and the Cambridge theories are not. This discussion has been to a certain extent misleading, because what makes a theory neo-classical or neo-Keynesian is whether or not it relies on factor scarcity and marginal productivities to explain the distribution of income. The question, whether a production function or some other concept—such as a technical progress function—is the correct way of representing the relation between factors and product in the presence of technical progress, must arise in any theory of growth.

It is clear that if the rates of change of two variables are related, their absolute values are also related, but the production function and the technical progress function are not interchangeable concepts. Suppose, in fact, that we had discovered the relation between the growth rates of output and capital *per capita* in a certain economy; in order to find out the relation between output and capital *per capita* at a certain date we would also need to know how long the growth process had been going on for and what the initial situation was. Furthermore, if we accept Kaldor's point that there is a ceiling to the flow of innovations per unit of time, so that the response of output to investment varies with the intensity with which the innovatory potential is exploited, we would need to know the precise pattern of investment over the past as well as the starting-point before we could determine the relation between output and capital *per capita* at a certain date. Consequently, we cannot abstract from history and define an aggregate production function as a purely technical relationship between output, labour and the stock of capital—a difficulty quite distinct from the 'aggregation problem', examined in the last section, of whether an aggregate production function exists, given the existence of individual production functions.

5. *Long-term economic policy in a neo-classical world*

In spite of its shortcomings, neo-classical growth theory has provided almost the sole logical basis for the theory of long-term economic policy as it has been developing in the last decade. We have already noted that the long-term rate of growth in output *per capita* cannot be a target of economic policy, unless the government has the means for influencing the rate of technical progress. Con-

sequently, the usual fiscal and monetary policies affecting the division of the national product between consumption and investment can only have transient effects on the *growth rate*, but their effects on the *absolute level* of output per efficiency unit of labour are, as we know, permanent.

We have seen that, if production and distribution take place according to the rules of the neo-classical model, an increase in the propensity to save of the community would lead to a permanent increase in capital and, therefore, output per efficiency unit of labour. But, while saving and investment per efficiency unit grow in proportion with capital per efficiency unit, output per efficiency unit grows less than in proportion because of the falling marginal productivity of capital. The difference between output and investment per efficiency unit—consumption per efficiency unit—will therefore increase up to the point where the increase in output is exactly matched by the increase in saving and investment; beyond that point an increase in the fraction of income saved would permanently decrease consumption per efficiency unit of labour. If an economy found itself above the point of maximum consumption, an obvious policy target would then be to increase present *and* future consumption by discouraging excessive saving—unless the community enjoys accumulating equipment for its own sake. Vice versa, if the economy were below the maximum consumption point, a policy target could be to increase consumption in the long term by increasing the fraction of income saved, though this would involve an immediate fall in consumption because output would take time to increase. The decision on how close to maximum permanent consumption to get would depend on the extent to which the community or its rulers are prepared to sacrifice the present generation for the next.

Some useful implications follow from the fact that consumption per efficiency unit of labour is maximized when a marginal increment in capital per efficiency unit of labour would increase output and investment per efficiency unit of labour by exactly the same amount. The rate of increase in output per efficiency unit of labour is in fact the marginal productivity of capital—assumed equal to the rate of return on capital or interest rate—and the rate of increase in investment per efficiency unit is the growth rate of the effective labour force or natural rate of growth: the rate of return on capital is therefore equal to the natural rate of growth of the economy. A further implication is that, as saving per efficiency unit of labour equals capital per efficiency unit times the rate of return on capital at the point of maximum consumption, saving is equal

to profits and consumption equal to wages. These two equivalent statements are known as the 'golden rule' of accumulation, first derived by Swan in 1960, and provide a clear guideline for long-term economic policy. A government that wanted to steer the economy towards the growth path on which consumption is at a maximum would then have to devise a tax system that gave maximum incentive to the consumption of wages and the reinvestment of profits; alternatively, it could aim its fiscal and monetary policies at keeping the market rate of interest as close as possible to the long-term rate of growth of the economy.

Whether or not governments want to maximize permanent consumption, it is now clear that it makes sense to talk of long-term policies. In order to evaluate the long-term effects of government intervention, we must bring public expenditure, taxation and the money supply explicitly into the analysis; this will also bring us a step closer to reality, and some of the conclusions reached on the basis of the cruder model examined in Section 3 above will need revising. A dynamic model with a government and money in it was presented by Tobin as early as 1955, and the need to take money into account was acknowledged by Solow in his article of 1956; more recently, the field has been further explored, mainly by Tobin himself, Johnson and Sidrausky—all American economists.

We start by observing that the government raises taxes and buys goods and services for public consumption from the private economy. But the government budget does not need to balance, because the deficit can be financed by issuing to the public a variety of debt instruments, some bearing interest, some not, depending on the time of maturity; the outstanding government debt is an asset in which the public can choose to put savings as an alternative or in addition to real capital. Let us call this composite paper asset 'money' for short, though it includes items not usually regarded as such; this is 'outside money' for the private economy, in the sense that, unlike the paper asset issued by the business sector, it is not matched by any private liability—there is no real capital behind it. Being perfectly divisible, easily transferable and guaranteed by the authority of the government, money is used as a convenient medium of exchange as well as a store of savings; consequently, there are now absolute as well as relative prices to contend with, and quantities can be measured in monetary as well as in 'real' or physical units. The supply of money will increase or decrease according to whether the government has a deficit or a surplus.

For full employment, the amount of output that can be efficiently produced with the labour and capital available at any point in time

must always be matched by the total demand for private and public consumption, and for investment. But real consumption cannot depend on real income alone as in the simple neo-classical model; it will depend on the rate at which income is taxed, on the stock of money held by the consumers and on changes in the purchasing power of money; that is, on the rate of inflation. Let us say that the community wishes to consume a constant fraction of its 'real disposable income', defined as real income after taxes, minus any losses on money holdings due to inflation. The level of investment demand required for any given level of output is then not fixed, but it varies in direct proportion to the rate of growth in the money supply, and in inverse proportion to the rate of inflation and the level of public expenditure. Consequently, the warranted rate of growth in capital does not depend on the propensity to save and capital/output ratio only, and the steady-state ratio of capital to output cannot be derived from the equation of the warranted and natural rates of growth unless we know the rate of inflation, the rate of growth in the money supply and the rate of public expenditure.

Public expenditure and money supply are under the direct control of the government and can be treated as policy instruments, but not prices; so we need additional information on the preferences of the community before we can find the growth path of the economy. What we need to know is how the community chooses to split its savings (or 'wealth') between capital and money. This decision depends essentially on a comparison between the rates of return on the two assets: the real rate of return on capital, equal to the marginal productivity of capital, and the real rate of return on money, which is the difference between the average interest yield of the individual assets and the rate of inflation. Other things being equal, the community would want to hold more money if the government increased the interest rate on the public debt, and less money if inflation accelerated or if there were an increase in the marginal productivity of capital. But marginal productivity falls with a more intensive use of the factor, and vice versa; therefore, reasoning in terms of ratios, we can say that the desired ratio between the stock of money and the flow of monetary income—i.e., the inverse of the 'velocity of circulation' of money—is directly related to the rate of inflation and the capital/output ratio.

Now suppose that the government has fixed the rate of public expenditure, the interest rate on the public debt and the rate of growth in the money supply. This leaves three unknowns—rate of inflation, velocity of circulation of money and capital/output ratio

—subject to three restrictions. One is the Harrod-Domar condition that the natural rate of growth must be matched by the warranted rate of growth, which we know to depend on the capital/output ratio and the rate of inflation. Another restriction is that the ratio between money stock and monetary income preferred by the community, which we said depends on all three of our unknowns, must be equal to the ratio determined by government policy. The third restriction is found by reflecting on the fact that the proportional rate of change in the velocity of circulation of money is given by the rate of growth in output, plus the rate of inflation, minus the rate of growth in the money supply. But, in a steady state, the rate of inflation must remain constant to allow a constant warranted rate of growth and, as the capital/output ratio and the rate of interest on money are also constant, velocity will have to be constant too. The rate of change in velocity is therefore zero and the steady-state rate of inflation must be equal to the difference between the rate of growth in the money supply and the natural rate of growth in output.

We can now see that government decisions determine the steady-state behaviour of the economy: the rate of inflation is the difference between rate of growth in money supply and natural rate of growth; this rate of inflation combined with the rate of public expenditure determines a capital/output ratio that will make the warranted rate of growth equal to the natural one; the rate of interest on money, together with the rate of inflation and the capital/output ratio, determines a velocity of circulation of money that will make the desired stock of money equal to supply, the capital/output ratio then determines output, consumption, wages and profits according to the usual neo-classical rules. It can be shown that—under neo-classical assumptions about the relationship between product and factors, and analogous assumptions about the relationship between demand for money and rates of return on the assets—there can only be *one* combination of the three unknowns which is consistent with a given choice of the three policy instruments in the hands of the government.

Once a mix of fiscal and monetary policies has been chosen, the course of the economy is thus set—but it cannot be shown that this growth path is *stable*, that is a monetary economy may not automatically tend towards it. Before money and absolute prices were taken into account we found that, whenever the warranted rate of growth diverged from the natural one, the factor mix and therefore the capital/output ratio were automatically adjusted; the economy then moved back to a steady state. But we have just seen that in a

monetary economy the capital/output ratio is tied up with the rate of inflation: if the capital/output ratio were, say, too high, there would be an excess of demand for both the composite commodity and money. The rate of inflation would then increase and there is nothing in what we have said so far to guarantee that it would go back to normal even if the capital/output ratio were eventually corrected—on the contrary, it is likely that inflation would feed on itself, perpetuating the imbalance. The fact that, in reality, economies tend to oscillate around a path of steady growth and not to run away from it must either mean that something essential is missing from the theory, or that a monetary economy is really unstable, and it is the frequent adjustment of short-term economic policy that keeps it in the vicinity of a golden age.

We can now answer the question of how changes in economic policy would affect the economy in the long term though keeping always in mind that the economy is likely to be unstable. There is, in fact, no guarantee that the economy will automatically move towards a new steady state once a change in the policy mix has thrown it off course—it may have to be guided towards the new golden age through a series of short-term policy measures, which in practice may prove very difficult to achieve. The targets of long-term policy are two: one is the capital/output ratio, because, the higher this ratio, the higher are output, wages and, up to a point, consumption per efficiency unit of labour—which given full employment means higher consumption per head of population; the other is the rate of inflation, because balance of trade considerations may not allow prices to grow faster than in the rest of the world, and because inflation is hard on fixed income earners.

If the government wanted no inflation at all, it could increase the money supply at the natural rate of growth of output—but the corresponding level of consumption per head might be far short of the maximum indicated by the golden rule. The government might then be prepared to put up with some inflation for the sake of permanently higher consumption. Inflation, in fact, reduces the purchasing power of the stock of money held by consumers, thereby increasing the fraction saved out of total income, and at the same time reduces the real rate of return on money, making investment in real capital more attractive. By accelerating the supply of money, the government would therefore increase not only the rate of inflation, but also the capital/output ratio and consumption per head. This leads to two considerations. The first is that money is *not* 'neutral'—in the sense of limiting its effects to the price level—because in the long term it affects real variables such as capital and

consumption. The second is that the more money the government puts into the economy, the less of it people want to hold (in proportion to income)—and the better off everyone would be if inflation did not matter.

A straight choice between high consumption and little inflation would be avoided if the government fixed the rate of increase in the supply of money so as to give an acceptable rate of inflation, and reduced instead the interest rate on the public debt so as to encourage people to hold more real capital relative to money. But, whichever target the government has in mind when fixing the rate of growth in the money stock, this decision will condition the choice of fiscal policy: the rate of increase in the public debt will in fact dictate the size of the government deficit (as a proportion of the national product) and the only choice left is between high or low rates of taxation and public expenditure, which will depend on the government preferences between private and public consumption. Consumption *per capita* and the rate of inflation can therefore be controlled, while keeping the economy along a path of full employment growth, only through concerted monetary and fiscal action.

The introduction of money and a government sector makes the neo-classical model much richer, and allows some Keynesian elements to be incorporated in the analysis—in much the same way as Keynes' short-term analysis was absorbed into neo-classical equilibrium theory in the hands of Hicks and Samuelson. In many ways this takes the neo-classical analysis further than that of the Cambridge school; but the basic objections to the neo-classical approach —that it does not recognize the independent role of the business sector in the decision to invest and to innovate, and that it relies too heavily on the assumptions of a well-behaved production function and perfect competition—are not yet answered.

BIBLIOGRAPHICAL NOTE

Much of the literature on growth theory is highly technical in nature and presented in mathematical form; the non-mathematical reader would be well advised to approach the subject via some good reviews. The one by F. H. Hahn and R. C. O. Matthews, 'The theory of economic growth' (*Economic Journal*, 1964), covering both the Cambridge and the neo-classical approach but without the monetary developments, is now a classic; a more up-to-date review of neo-classical theory alone is in R. M. Solow, *Growth theory*

(Oxford University Press, 1970). A good but advanced textbook on neo-classical growth models is E. Burmeister and A. R. Dobell, *Mathematical theories of economic growth* (Macmillan, 1970); the Cambridge and the neo-classical approach are discussed less technically, with emphasis on the underlying theories of distribution, in J. A. Kregel, *Rate of profit, distribution and growth* (Macmillan, 1971).

The original references on the Harrod-Domar model are: R. F. Harrod, 'An essay on dynamic theory' (*Economic Journal*, 1939) and *Towards a dynamic economics* (Macmillan, 1948); E. D. Domar, 'Capital expansion, rate of growth and employment' (*Econometrica*, 1946), reprinted with other important essays in *Essays in the theory of economic growth* (Oxford University Press, 1957). On technical progress see J. Robinson, 'The classification of inventions' (*Review of Economic Studies*, 1938), and H. Uzawa, 'Neutral inventions and the stability of growth and equilibrium' (*Review of Economic Studies*, 1961).

On the basic neo-classical model see R. M. Solow, 'A contribution to the theory of economic growth' (*Quarterly Journal of Economics*, 1956), and T. Swan, 'Economic growth and capital accumulation' (*Economic Record*, 1956). On the aggregate production function see R. M. Solow, 'Technical change and the aggregate production function' (*Review of Economics and Statistics*, 1957), and F. M. Fisher's excellent review of the aggregation problem, 'The existence of aggregate production functions' (*Econometrica*, 1969). The whole debate on capital and the production function is well summed up by G. C. Harcourt in 'Some Cambridge controversies in the theory of capital' (*Journal of Economic Literature*, 1969), which contains an exhaustive bibliography.

Kaldor's theory is in 'A model of economic growth' (*Economic Journal*, 1957); see also his 'Alternative theories of distribution' (*Review of Economic Studies*, 1955) and 'Capital accumulation and economic growth' in *The theory of capital*, edited by F. A. Lutz and D. C. Hague (Macmillan, 1961), and L. Pasinetti's 'Rate of profit and income distribution in relation to the rate of economic growth' (*Review of Economic Studies*, 1961). The 'vintage' version of the model is in N. Kaldor and J. A. Mirrlees, 'A new model of economic growth' (*Review of Economic Studies*, 1961). On the comparison between production and technical progress functions see J. Black, 'The technical progress function and the production function' (*Economica*, 1962), and F. H. Hahn and R. C. O. Matthews, *op. cit.* A similar model, in which technical progress is explicitly seen as the result of cumulative experience, is that of K. J.

Arrow, 'The economic implications of learning by doing' (*Review of Economic Studies*, 1967).

On consumption maximization see E. S. Phelps, 'The golden rule of accumulation: a fable for growthmen' (*American Economic Review*, 1961), and I. F. Pearce's critique in 'The end of golden age in Solovia' (*American Economic Review*, 1962). On the introduction of money and public spending in a growth model see J. Tobin, 'A dynamic aggregative model' (*Journal of Political Economy*, 1955), and 'Money and economic growth' (*Econometrica*, 1965); H. G. Johnson, 'The neo-classical one-sector growth model' (*Economica*, 1966); M. Sidrausky, 'Inflation and economic growth' (*Journal of Political Economy*, 1967), and 'Rational choice and patterns of growth' (*Journal of Political Economy*, 1969).

Chapter 10

THE THEORY OF ECONOMIC PLANNING

1. *The origins of the debate*

In the course of this book we have often referred to the theoretical problems of a *planned economy*: it is now time to deal explicitly with this subject, which occupies an important place in the history of modern economic thought. By planned economy we mean an economy in which all the important economic decisions are taken by the government. We are therefore excluding all those forms of public intervention in market economies that are sometimes referred to as 'planning'.

The origin of the controversy on the theory of economic planning can be found in the position taken by the Austrian economist Von Mises in an article published in German in 1920 and better known in its English translation of 1935, in which he denied that a planned economy could ever work rationally. It may seem strange that the beginning of the discussion should be traced back to an opponent of economic planning, instead of looking for its origins in the Marxist literature. The point is that both Marx and Engels had anticipated the need for a planned management of production after the Socialist Revolution, but neither of them had given any thought to the way in which a planned economy would have to function: Marx's refusal to provide recipes for the future is well known. As for the Soviet economic literature, something interesting started to emerge only well after the time when Von Mises started the debate. The reasons for starting this exposition with the Austrian economist are therefore chronological, but we will see that there are also logical reasons.

2. *The criticisms of Von Mises, Hayek and Robbins*

Von Mises' argument for accusing the planned economy of irration-

ality is the following: the purpose of any economy, whatever its institutional framework, is that of utilizing scarce resources in the most efficient way towards the achievement of certain ends. But, in order to use resources efficiently, there must be precise indexes of scarcity, in the absence of which the choices made would not be efficient. Now the only known method of attaching scarcity indexes to the productive resources is to allow the resources or their services to be exchanged on the market, for their relative prices will reflect their relative scarcity. These prices are therefore a necessary starting-point for making rational choices but, as a planned economy has by definition no markets, it has no mechanism for measuring the relative scarcity of resources, no rational criterion of choice, and is therefore doomed to chaos. A planned economy can, according to Von Mises, produce outstanding material results, but at what cost no one will ever know. As a matter of fact, as the cost cannot be calculated, it is only by sheer chance that a planned economy can avoid waste on a gigantic scale.

A reply *ante-litteram* to Von Mises' criticism had already been given by Enrico Barone (in the article already quoted in Chapter 2 above) and was implicit anyway in the whole theory of general economic equilibrium. In fact, if one accepts that the mechanism which determines the choices of the individuals, i.e., the mechanism that generates relative prices, can be put in the form of a system of equations, then one has to accept, at least in principle, that given the terms of the problem the prices can be calculated without any need for a market. The terms of the problem that have to be known are the amounts of resources and technology available and the preferences of the consumers. If these data can be measured statistically the general equilibrium system will tell us what is the most efficient configuration. This point was realized by many economists who, nevertheless, followed Von Mises in denying the possibility of a rational planned economy. The most important among them are Hayek and Robbins who, unlike Von Mises, accepted that in principle an economy could be rationally planned, but denied that this could be done in practice for two reasons, Firstly, because the general equilibrium system of equations is so large and complex that there can be no hope of actually finding a numerical solution to it. Secondly, because even in the hypothesis that the system could be solved the calculation would require such a long time that the statistical information on which the calculation would be based would be out of date by the time the results were found.

The position of these economists is therefore much less rigid than that of Von Mises: for Von Mises a planned economy is in-

conceivable; it is not even a proper economy but only an uncoordinated set of activities. For Hayek and Robbins, on the other hand, a planned economy is quite conceivable in theory but could not be put into practice.

3. *The reply of the Marxists: Lange and Dobb*

What was the reaction of the supporters of economic planning? It must be said straight away that there was no united reply and that the main contributions to the discussion have to be examined separately. The first chronologically was that of Lange, in an article published for the first time in English in 1936. Lange's thesis has to be looked at carefully, because it has provided one of the starting-points for modern discussion of economic planning. Lange starts by rejecting Von Mises' position on the basis of the general equilibrium arguments already mentioned. He argues that once it is recognized that a price is not merely a rate of exchange, but also an index of scarcity, the idea of a planned economy can no longer be rejected in principle. The concept of price as an index of scarcity makes it in fact possible for it to be calculated even outside a market system, that is, whether or not an act of exchange has actually taken place. For Lange, the fact that prices thus conceived can be calculated without any reference to the market is a landmark in the history of modern economic thought that can and must be exploited for planning purposes. But, although opposed to Von Mises' position, Lange reveals himself very receptive to Hayek's and Robbins' arguments. In fact, he cannot imagine how relative prices would be determined in a planned economy, except through a mechanism which would be an exact replica of the competitive market. This seems to imply that Lange does not see the possibility of actually solving a general equilibrium system and ends up by recommending an institution very similar to a market.

The institution proposed by Lange is, however, very different from the capitalistic market. Firstly the new institution would consist of a set of publicly owned productive units each of which would be completely free to make its own decisions within certain rules of the game dictated by the authorities; these rules are none other than those one would expect to be followed by a profit-maximizing firm under conditions of perfect competition. Secondly, there must be a central authority with the function of fixing the 'accounting prices' on the basis of which the individual productive units will

make their decisions; these prices would have to be such that demand and supply are equated for each good, and would have to be arrived at through a process of trial and error. In Lange's view such a mechanism would be much more efficient than the capitalist market, because the latter is characterized by various monopolistic imperfections, while in a socialist economy, in which all means of production are publicly owned, it would be possible to set up an 'ideal' market.

Lange's position requires very careful consideration, because it contains both positive and negative elements. The main positive element is that for the first time Von Mises' criticism is taken seriously by a supporter of economic planning. Before Lange, in fact, the socialist literature had failed to realize the nature of the problems facing a planned economy. It had failed to recognize that the problem of rationality is not, as it was held, peculiar to the capitalistic economy, but a general problem, because any economic organization has to find a way of efficiently utilizing scarce resources. Lange's great merit is in having clarified this point and also in having shown the relevance of a whole body of economic theory of Walrasian extraction for solving rationally the problems of economic planning. The weakness of Lange's argument is in the fact that he resorts to a mechanism not dissimilar from a market to solve the problem of calculating the equilibrium prices. This would in fact waste some of the most important advantages of a centrally planned economy. One could even say that the socialist economy described by Lange could hardly be called a planned economy, because the decentralization of economic decisions is taken to the extreme. This would get rid of the main feature of a planned economy and, for this reason, many socialist economists have tended to move in a substantially different direction.

The main exponent of this different school of thought is the Englishman Maurice Dobb, who reaches conclusions different from those of Lange because he starts with a different criticism of Von Mises' arguments. The difference is that while Lange accepts the problems posed by Von Mises and tries to show that they can be solved in a socialist economy, Dobb does not accept the problem as being of crucial importance, and, by doing so, puts himself in the main stream of socialist tradition.

Dobb argues that Von Mises' problem concerns the efficient use of *given* resources, while economic planning is a powerful instrument for solving the problem of accumulation; that is, the *increase* in the amount of resources available. Dobb therefore accepts that economic calculations can run into very serious difficulties in a

planned economy, but he considers this problem as being of secondary importance. The effectiveness of central planning as an instrument of accumulation derives from the fact that planning is a substitute for the market. His thesis is that, while the market is a mechanism for co-ordinating the investment decisions *ex-post*, central planning makes it possible to co-ordinate them *ex-ante*. The market can only make reciprocally consistent the decisions already taken via movements in relative prices. In a planned economy, on the other hand, the co-ordinating mechanism is the central plan itself, which consists of a co-ordinated set of initiatives assigned by command to the economic system as a whole.

Dobb's view is that, in an age in which technical progress has led to a considerable increase in the size of the individual decisions to be taken, the *ex-ante* co-ordination provided by central planning is definitely superior to the *ex-post* co-ordinating mechanism provided by the market. In fact, the size of each individual investment decision is now such that its success depends on a number of other well-defined decisions being taken at the same time at other points of the economic system and, moreover, it does itself influence those other decisions. Besides, again as a consequence of the nature of technical progress, the quota of fixed investment in total investment tends to become larger and larger, so that more and more national resources tend to be committed to a particular use for long periods of time. Confronted with these circumstances, the market mechanism generally proves to be inadequate, as shown by the fact that many developments, which would have appeared perfectly feasible if there were full information about the economic system as a whole and its future expansion, are regarded as unfeasible (that is, non-profitable) by the individual firms. Particularly in those countries in which economic growth is the main target, central planning has a decisive advantage over the market. We could conclude by saying that, in countries that intend to initiate or to accelerate the process of economic growth, the market mechanism would attach such great risk to the most important development decisions that this would appear completely unprofitable; central planning would then be the inevitable choice.

The positions of Lange and Dobb can be regarded as being representative of two different ways of reacting to the criticism that no correct economic calculation is possible in a planned economy. The first type of reaction consists of three propositions: it is accepted that economic calculations are essential in a planned economy as in any other economic organization, and that no calculation is possible without an institution analogous to the market; but it is main-

tained that a mechanism analogous to the competitive market can be introduced in a planned economy and would be much more efficient than a capitalistic market. The second type of reaction, vice versa, consists of the two following propositions: it is denied that economic calculations are essential to economic planning, because the main problem is not that of optimally utilizing given resources, but that of increasing the amount of resources; and, for this purpose, the plan is more efficient than the market, for it makes possible developments that would be inconceivable in a market economy.

4. *Planning and economic efficiency*

We have already seen where the weakness of Lange's argument lies; but Dobb's argument also has a weak spot. The problem of efficiency does not in fact apply only to a static situation (although it is mainly in that context that it has been studied) but applies equally well to intertemporal choices. There is no doubt that the economic notion of efficiency—maximizing the achievement of certain ends with given means—implies the assumption of given productive resources; but one of the ends can be that of increasing the resources available, and this, broadly speaking, is what accumulation is about.

Accumulation can therefore be efficient or inefficient. The fact that part of the resources available are dedicated to increasing the productive capacity of the system does not avoid the problem of using those given resources efficiently. Remember, in fact, that accumulation is a goal conflicting with others aiming at the satisfaction of immediate consumption, and that even the aim of increasing productive capacity can be achieved in many different ways, between which one has to choose. Therefore, the choice between consumption and accumulation, and between the many alternative modes of accumulation, cannot be made without a well-defined criterion of choice and without some mechanism for finding the best alternative. This does not necessarily mean that the market works better than the plan even in a situation in which the problem of growth is the most important; it simply means that the superiority of central planning cannot be demonstrated unless there are adequate instruments for measuring the efficiency of a planned economy. In other words, economic calculations are necessary whenever there are scarce resources, and the fact that an

economy has set itself the target of making these resources less scarce through accumulation does not mean that resources are no longer scarce: the aim can be efficiently pursued only if there is a mechanism that allows economic calculations.

It is now clear that an adequate reply to the critics of economic planning should contain the positive elements of both Lange and Dobb and should be able to show that it is possible to have a form of organization that allows economic calculations but one which, at the same time, does not present the shortcomings of the market vis-à-vis the problem of growth. This is the direction in which the theory of economic planning has been developing recently. Before proceeding any further, however, we ought to have a quick look at the Soviet experience.

5. *The Soviet experience*

This exposition will have to be necessarily brief not only for lack of space, but also because this book is a history of ideas rather than facts. It is interesting to note that up until very recently Soviet planners have almost completely ignored the problem of economic efficiency. At the root of this attitude were two circumstances, one of a cultural and the other of a practical nature.

The cultural circumstance is that Soviet economists have never attempted to do more than illustrate the thoughts of Marx, who, as we know, had no detailed views on the kind of organization that was to replace capitalism. Even Marx's analysis of capitalism was primarily concerned with those aspects that are peculiar to that form of organization and, therefore, it is not possible to find in his works an adequate treatment of the problem of the efficient utilization of scarce resources, which is not, as such, a capitalist peculiarity but a common feature of all economic systems.

The practical reason for the Soviet attitude can be found in the fact that, at least in the first stages of economic planning, the alternatives were so straightforward that the choice could be made even without rigorous calculation. The choice was simple because, firstly, the dilemma between present and future consumption had been resolved by the decision to have the highest possible rate of economic growth. Secondly, there was the fact that at the beginning of a process of industrialization there are only a few simple ways in which capital can be accumulated. Thirdly, even the choice of what to consume was simple because, having decided to keep con-

sumption to a minimum for the sake of accumulation, the consumer goods to be produced were only the bare essentials. In such circumstances the only problem to worry about was that of ensuring that the developments undertaken in the various sectors of the economy were reciprocally consistent, so as to ensure that the demand and supply of each good were in balance at each point in time. This problem was solved by the method of the so-called 'material balances', which is a sort of rudimentary Leontief system. But, even if this method of 'material balances' had been developed into a proper input-output model, it would still have guaranteed only the internal consistency of each possible growth alternative; it could not have solved the problem of the choice between several growth alternatives.

As a matter of fact, the pre-1927 Soviet economic literature contained some discussion of the problem of economic efficiency, and there were some important contributions which, however, in the cultural and political climate of the time had no influence whatever on the policy-makers and later ceased altogether. On the other hand, as the process of industrialization proceeded and prosperity increased, the planned management of the economy became more and more complicated as the choices to be made became less and less obvious. On the one hand, the objective of rapid industrialization made it necessary to choose between alternative and increasingly diverse industrial structures; on the other, the increase in the general standard of living, achieved in spite of the high rate of accumulation, made it possible to choose from a much wider range of consumption alternatives. The problem of choice and therefore of efficiency could no longer be ignored.

A very interesting indication of this situation can be found in the last of Stalin's writings, published in 1952. Stalin points out that there are two areas within the planned Soviet economy in which prices are still determined by market forces. These are the market for consumer goods and the market where farm produce is exchanged for industrial goods. In both markets we meet independent economic operators. On the consumer good market the mass of the consumers, left free to spend their income as they please, meet the whole of the government retail organizations. On the other market the State industries meet the collective farms. Within certain limits these are both proper markets in which prices are formed through bargaining.

The interesting point about Stalin's position is that he felt a hurried elimination of these market forms from the Soviet economy would be a mistake. On the contrary, his opinion was that these

markets contributed to the rational management of the economy—an objective to which he attached great importance, because of the waste and contradictions present in the Soviet economy, of which he gives some examples himself in this essay. Stalin's approach to the problem of rationality is therefore still based on the idea of a market, and the problem of how to manage the Soviet economy is solved by a mere compromise between the centralized planning of long-term growth and the decentralized determination of prices in some sectors.

Stalin's position certainly reflected widespread scientific opinion in the Soviet Union at a time when attempts were being made to put the whole planning mechanism on a more solid footing, under pressure from the innumerable problems emerging from the practice of planning. A problem that had particularly interested both planners and economists for a long time was the criterion for choosing between different modes of carrying out an investment (for example, whether to increase the output of electric power by building additional thermic or hydro-electric plants). It was realized that it was not even possible to compare investment alternatives without discounting to present time the future costs and revenues expected from each of them, and that this operation was not possible without first determining an interest rate. To admit even the existence of a rate of interest was an important departure from Marxist orthodoxy, which regarded the interest as a consequence of capitalistic exploitation, and the discussion went so far as to debate whether the rate of interest should be the same for the whole economy or vary from one project to another. So Soviet planning was faced with the same efficiency problems that face a capitalistic economy and it is therefore natural that the importance of an economic category such as the interest rate, which is essential for the solution of these problems, should at last be recognized.

We could conclude by saying that the level of economic development reached by the Soviet Union has made it impossible to go in with the approximative methods of management of the first stage of industrialization: the waste of resources arising from the lack of rigorous methods has started to become unacceptable. In addition to this, there is the unhappy planning experience of the other Eastern European countries, where the ill effects of the lack of rigorous economic criteria are so conspicuous that some Polish economists have come to state that the low standard of living in their country is due much less to the high rate of accumulation than to the waste of resources deriving from inefficiency.

The present situation can be described as follows: the level of economic development and industrial sophistication reached by planned economies is now such that the problem of managing these economies efficiently can no longer be postponed. Although it is generally recognized that one has to be very careful before completely eliminating what is left of the market mechanism within a planned economy, nevertheless the idea that calculations for the economy as a whole can be based on such mechanisms tends to be ruled out, because of the advantages of central planning from the point of view of long-term growth. The problem is to see whether the central planner can carry out all the necessary calculations without relying on the automatic mechanisms of the market, whether this idea which has seemed absurd to Hayek and Robbins, and in their wake to many friends and enemies of central planning, has now become a practical proposition.

What has changed today is that the planners have at their disposal a whole array of mathematical models capable of reproducing the mechanism of a perfectly competitive market, if ever there was one. More so, these models have now been made operational, that is, they are designed in such a way that with the help of modern computers they are capable of giving actual numerical solutions. We have in fact seen, in the first chapter of this book, how the analogy between the competitive mechanism and planning had already been stated by Pareto and, more explicitly, by Barone. But an actual plan could never be built on the basis of the Walrasian theory, because of the shortcomings of this theory in the area of capital formation. The most accomplished planning models now available are therefore based on Leontief's dynamic model, mentioned in Chapter 8.

Because of the phenomenon of 'duality', mentioned in Chapter 8, there are two possible alternative approaches to economic planning, according to whether the central planners dictate the actual productive targets or determine instead a set of prices. In the first case the decision-making process is completely centralized, because the central planners assign specific and detailed production targets to the individual productive units. In the second case there is a strong element of decentralization, because the planning authorities decide only the price system corresponding to the desired productive configuration, but the actual production decisions are taken, on the basis of those prices, by the individual productive units. In practice, it is possible to find many intermediate situations

where the planners find it expedient to dictate productive targets in certain cases and to fix prices in others.

Some recent trends in the economic thinking of the Soviet Union seem to indicate that that country is now definitely orientated towards the application of the more sophisticated planning methods that we have been talking about. In the case of some operational models now in use in the West, the Soviet Union can even boast a priority of which it has only now started to appreciate the value. The method of linear programming, for example, was formulated for the first time by the Russian mathematician Kantorovitch in 1939. The same Kantorovitch, together with numerous other Soviet economists, has been working on the construction of the tools for a more rational management of the Soviet economy. It is also interesting to examine the later position of Lange, who, after returning to Poland from the United States in 1945, occupied positions of responsibility in Polish planning organizations until his death in 1965. In his later writings Lange detaches himself quite radically from the position held in 1936 and begins to elaborate planning schemes founded on centralized decision-making which rely on mathematical calculation (rather than on the *tâtonnement* procedure of organizations analogous to the market) for finding efficient solutions.

We are obviously at the beginning of a phase of change in the theory and practice of economic planning and it is difficult to predict with any certainty what future developments will be because the planning instruments that have been elaborated so far, although infinitely more refined and effective than those which had been used before, still have many defects. One of the most serious defects seems to be the very unsatisfactory way in which technical progress and all the related phenomena are dealt with. It is, nevertheless, a very important turning-point which will have long lasting effects not only on the practice of economic planning but also on the development of economic thought.

BIBLIOGRAPHICAL NOTE

The debate on the theory of economic planning starts with L. Von Mises' article 'Die Wirtschaftsrechnung in Sozialistischen Gemeinwesen' (*Archiv für Sozialwissenschaften*, 1920); an English translation of this article and of Barone's *Ministro della produzione* are included in *Collectivist economic planning*, edited by F. A. Hayek

(Routledge, 1935). In the same volume, see also Von Hayek's contributions, 'The nature and history of the problem' and 'The present state of the debate'. Robbins' views are in *The Great Depression* (Macmillan, 1934).

Oskar Lange's initial position is in the article 'On the economic theory of socialism' (*Review of Economic Studies*, 1936 and 1937), reprinted with some modifications in the book '*On the economic theory of socialism*, edited by B. E. Lippincott (Minnesota University Press, 1938, 2nd edition 1948). The subsequent developments of his ideas are reflected in a long series of works published since 1958. See, in particular, *Wstep do ekonometrii* (1958), translated in English as *Introduction to econometrics* (Pergamon, 1959, 2nd expanded edition 1962)—the first systematic treatment of the use of mathematical and statistical methods in a planned economy—and, in Italian, 'Il ruolo della matematica nella pianificazione economica' (*Statistica*, 1963).

By M. Dobb, see *On economic theory and socialism* (Routledge, 1955) and *An essay on economic growth and planning* (Routledge, 1960).

On the problems and debates of Soviet Russia, see J. V. Stalin *Ekonomischeskih problemy Sotsializma VSSSR* (1952), translated as *Economic problems of socialism in the U.S.S.R.* (Foreign Languages Publishing House, Moscow, 1965), and L. V. Kantorovich, *Ekonomicheski raschet nailuchsheqo ispol 'zovaniya resursov*, edited by V. S. Nemichinov (1959), translated as *The best use of economic resources* (Pergamon, 1965, 2nd edition 1968).

Chapter 11

THE THEORY OF ECONOMIC UNDER-DEVELOPMENT

1. *The roots of the problem*

During the last twenty-five years economists have been applying their minds to an economic reality that in the past had received only scant attention: the reality of all those countries that are now known as the under-developed world. The study of the subject started during the last World War, when it was realized that at the end of the conflict many countries of Eastern Europe, Asia, Africa and Latin America would have posed the problem of how to narrow the gap between themselves and the economies of the more advanced countries. It was in fact realized that the sense of national identity, enhanced by the war itself, made it impossible for the poorer countries to accept any longer the enormous imbalance in the world distribution of material wealth.

Since then the literature on under-developed economies has grown to imposing proportions. This literature deals with three broad questions: the definition of under-development, the fundamental characteristics of under-developed economies and the way in which a situation of under-development can be overcome, that is the policies aimed at determining or facilitating a process of rapid economic development.

2. *The definition of an under-developed economy*

To give a definition of under-developed economy is not as simple as simple as one might think. As a matter of fact economists have not been able to agree on which of the many outstanding features of actual under-developed economies should be regarded as *the* characteristic of under-development. Nevertheless, it is common practice to use the level of income *per capita* of an economy as an index of its degree of development, for no better reason than that

it provides a handy numerical value. The shortcomings of the income *per capita* criterion have been amply illustrated in the literature and can be briefly summarized as follows. Firstly, there is the practical difficulty that under-developed countries are very poor as regards statistics on income and production, so that it is often impossible to obtain estimates of national income, and international or intertemporal comparisons based on this criterion are therefore of doubtful value. Secondly, there is the problem that any estimate of national income, however accurate, generally reflects only that part of the national product that passes through the market and can consequently be evaluated at market prices. But in under-developed economies a very large part of the national product does not go through the market at all, being the output of producers living merely at subsistence level who neither buy means of production nor sell their product. This means that there is no record of a considerable part of the national income of such countries. Thirdly, there is the difficulty of using income *per capita* for international comparisons, because of the serious problem of reducing national data to the same monetary unit. The rates of exchange between the various national currencies, in fact, hardly ever reflect their relative purchasing powers. On the other hand, if one attempts to calculate the ratio between the purchasing powers of two currencies, one finds that, when comparing a developed to an under-developed country, the 'baskets' of commodities consumed by the two countries are so different that even the concept of a ratio between purchasing powers becomes very dubious. Finally, there is the devious problem that a comparison based on income *per capita* completely overlooks any qualitative differences, as two countries with the same level of income *per capita* can have even very different economic structures.

In spite of these shortcomings, the criterion of income *per capita* gives a pretty good idea of the extent of imbalance in the world economy. We go, in fact, from incomes *per capita* of around $2,000 *per annum* (as in the United States) to less than $50 *per annum* (as in many African and Asian countries). The picture would certainly not change substantially if one could devise a method of measurement that met all the criticisms mentioned above. If one regards as 'under-developed' all the countries with an income *per capita* of less than $200 *per annum*, then 70 per cent of the world population lives in conditions of under-development. This datum alone, even taking into account all the imperfections of the criterion used, is sufficient to give an idea of the proportions of the problem.

3. *The economic structure of under-developed countries*

The immediate cause of low income *per capita* in under-developed economies is the low productivity of labour, which, in turn, is the result of the very primitive equipment with which labour is combined in the productive process. In order to escape from this situation of technical backwardness it would be necessary to accumulate capital fast and, in order to be able to do that it would be necessary to have a level of production in excess of the level of consumption required for the subsistence of the labour force. This is, in fact, what has happened in all the countries that are now developed. But because of the very low productivity of labour the excess of production over consumption is very small in under-developed countries, which therefore find themselves in a kind of vicious circle, the so-called 'vicious circle of poverty': to increase income *per capita* it would be necessary to increase productivity, which in turn would require a fast rate of accumulation of capital and therefore a substantial excess of production over subsistence consumption; but this surplus is very small because income *per capita* is so low in the first place. The existence of this vicious circle condemns under-developed countries to a practically stationary situation, which is all the more hopeless because the excess of output over subsistence consumption is not only small but, in many cases, also destined for the 'affluent consumption' of the wealthy classes.

At the root of this systematic tendency towards a stationary situation is the fact that under-developed economies are characterized by structures and forms of organization that can be defined as *pre-capitalistic*. What their economic structure lacks in fact are the two fundamental characteristics of capitalism: a labour force consisting of hired workers, and investment decisions taken by a separate category of people on the basis of their profit expectations. The capitalistic mode of production brings about a great extension of trade; firstly because no individual producer is self-sufficient, and, secondly, because of the continuous stimulus to invest in order to obtain and maintain a profit in the face of competition. This stimulus to invest is the powerful spring that has allowed the developed countries to accumulate in little more than a century an enormous stock of capital and therefore to achieve a rapid and systematic increase in the level of income and standard of living, through the continuous increase in the productivity of labour.

Another important difference between the economic structure of the developed countries and that of the underdeveloped ones is that in the former most of the output is the result of industrial produc-

tion, so that they are also referred to as the *industrialized* countries. In contrast, most of the output in the under-developed economies is the result of subsistence activities conducted on a family basis. This applies not only to agriculture, but also to manufacturing, which takes place mainly in family-sized, cottage-type rather than properly industrial productive units. But the prevalent economic activity of under-developed countries is still agriculture, generally employing between 70 and 90 per cent of the working population.

The pre-capitalistic nature of most economic activities explains the phenomenon of *concealed unemployment* which is typical of under-developed countries. By concealed unemployment we mean a situation in which the number of workers employed could be reduced without causing a fall in production, even without a change in the stock of capital and technique of production used. It is clear that a situation like this could occur only in an economy in which there is no net distinction between firm and family, as for example in a peasant agriculture. Only in such a case would it in fact be possible to find workers with zero productivity permanently employed in a firm: in a capitalistic organization such workers would soon be made redundant. On the other hand, the presence of concealed unemployment constitutes a reservoir of productive potential that, under certain conditions, would make it possible to initiate a development process. We shall come back to this point later.

The fact that what little is left of the national product after providing for the subsistence of the workers is destined for the affluent consumption of the wealthy class can be traced back to the pre-capitalistic nature of under-developed economies. The wealthy are, in fact, mainly landowners rather than industrial entrepreneurs. So the presence of a wealthy class, which can be a dynamic element in a capitalistic economy, is instead one of the strongest obstacles to development in backward countries.

4. *International disparities*

The great differences in economic structure between different countries help to explain the great inequalities in the international distribution of income. It is interesting to examine the prediction on international economic development made by the classical economists, on the basis of the capitalistic reality that surrounded them. Confronted with the obvious expansionary thrust of the capitalistic mode of production, the classical economists predicted that capital-

ism would completely replace all other modes of production in the countries in which it had originated; that is, in the countries that had experienced an industrial revolution. Capitalism would then expand to the rest of the world because the capitalist countries would be in a position to win and subject all others. Now, more than a century after these predictions were made, we find that they have come true only to a very limited extent. We find, in fact, that the countries that were already developed at the time when these predictions were made have become richer and richer, while most of the other countries have failed to enter the capitalist camp and are lagging further and further behind. As Myrdal explained very clearly, international economic history shows a tendency not towards international equilibrium, but on the contrary towards increasing disequilibrium.

Many reasons have been put forward as to why the classical predictions went wrong. One of those most generally accepted by the theorists of economic under-development is connected with the presence of *conglomerative factors.* These factors are such that, once there is an imbalance, however slight, between different regions, they tend to perpetuate and increase it, whatever the initial cause. These factors fall into two categories, depending on whether they act on the supply or on the demand side.

Conglomerative factors of the first type are in operation wherever industry happened to settle in the first place and gave rise to a series of 'external economies' for all firms that might choose to establish themselves in that area. These economies are the result of all the facilities (power, communications, credit, trained labour, etc.) available in an already industrialized area, which afforded a decisive advantage to all firms in that area. The classical economists thought that these advantages enjoyed by countries which were already industrialized would have been offset by the low cost of labour in the rest of the world. But historical experience has shown that this offsetting has not taken place: in the last century an Indian worker could be hired for one-fifth or one-sixth of the wage of an English worker, but nevertheless business always found it more profitable to invest in England than in India, because the external economies existing in England were much greater than the saving in the cost of labour that could be made in India.

The existence of conglomerative factors operating on the demand side can be best explained with an example. Imagine planning the installation in an under-developed area of a factory that will employ a thousand workers, previously unemployed. The new income earned by these workers obviously cannot be spent entirely

on the products of the new factory, which will therefore have the problem of finding other markets. The problem could be solved by employing the thousand workers not in one factory, but in a complex of smaller factories producing such a wide range of goods that the entire production could be absorbed by the workers themselves. But for modern factories there is a minimum size below which they cannot be operated, so that the only way of solving the problem would be to employ not a thousand but, say, ten thousand workers, creating a set of industrial units, each sufficiently large, producing a set of commodities sufficiently differentiated to satisfy the demand of the additional workers employed. So the problem of demand can be seen as a problem of scale. This implies that the existence of a certain degree of industrialization, and therefore of a certain number of productive units already in operation, gives an area a decisive advantage from the demand site too.

These arguments are summed up by Rosenstein-Rodan by saying that in under-developed countries there are problems of 'indivisibility', both from the side of demand and from the side of supply, in the sense that an investment project appears to be profitable only if it is larger than a certain minimum size. Consequently, a development process starting under modern technological conditions requires a very large initial investment in order to create, all at once, all those conditions that make an investment profitable. It is precisely for this reason that under-developed countries find it so difficult to get the development process off the ground and remain trapped in the vicious circle of poverty.

The existence of conglomerative factors is certainly important, but cannot explain the phenomenon of under-development completely. The argument based on conglomerative factors starts, in fact, from the consideration that if there is an initial imbalance, however slight, there is a mechanism that makes the imbalance itself larger and larger; but it does not explain how the initial disparity came about. The point is that the original imbalance cannot be explained in purely economic terms, because one has to consider all the historical reasons for which an entrepreneurial middle class developed in some places rather than others. The reasons are obviously not only economic, but depend on the culture, religion and pattern of civilization in general. All we can say about these other reasons is that they are very important and have to be kept in mind.

There is, finally, an important political aspect which needs to be clarified. We know that most under-developed countries are or have been colonies of industrialized countries. As part of the policy of

colonization, the industrialized countries made considerable investments in their colonies, aimed mainly at the exploitation of natural resources such as oil, rubber, copper, coffee, cotton, and so on. There is no doubt that these investments gave rise to technologically advanced and unmistakably capitalistic productive activities. But they could not change in any way the general picture of underdevelopment, because the object of these industrial activities was only that of making certain goods available to the industrial apparatus of the colonizing country and not to start off a self-sustaining development process in the colony. This is not to deny that industrial investment in the colonies has had any positive effect, because it has at least created a sector, however small, of modern economic activity which, by recruiting local manpower, initiated the destruction of the more primitive forms of productive organization. The ensuing crisis of the traditional section of the economy has been one of the reasons for which ex-colonial countries have had to look for new patterns of industrial development in which local resources could be better used to increase the standard of living of the population.

5. *Industrialization and development planning*

The development of a backward economy is generally seen in terms of its *industrialization*; in terms, that is, of the expansion of the activities with the highest productivity of labour, which would give the maximum contribution to increasing income *per capita.* But the problem has not always been seen in these terms. A superficial interpretation of the classical theory of international trade, which advocates the international distribution of labour, could in fact be used as an argument against industrializing the under-developed countries and for having a division of the world economy into agricultural and industrial countries. This view has now been rejected on the grounds that the classical thesis of the international division of labour starts from the assumption that the international distribution of productive resources is given and cannot be modified, while the aim of industrializing under-developed countries is precisely that of modifying such distribution. The process of industrialization, by changing the international distribution of capital, changes in fact the terms of the problem to which international specialization into industrial and agricultural countries was supposed to be the best solution.

It is therefore widely agreed that in order to develop economically a country has to industrialize, but the institutional framework within which the process of industrialization should take place is still the subject of contention. It is in fact clear that this process could hardly be left to the market mechanism prevailing in England and the other mature economies at the time of their industrial revolution. In those countries industrialization was a gradual process, in which the productive system developed bit by bit, progressively increasing its complexity. Meanwhile, the social infrastructure (transport, communications, public administration) had time to grow at the same pace, thus making possible and reinforcing the industrial growth itself. But this process could not take place in today's under-developed countries. Here it is a matter of making good lost ground, of reaching, within a reasonable period of time, a situation roughly comparable with that of the developed countries.

At this point it is useful to mention a problem which has been discussed at length in the literature on under-developed countries; a problem that has often been presented in a very complicated way, but is basically very simple. At the beginning of a development process an economy has a great abundance of labour relative to capital. This being the case, some economists, such as Ragnar Nurkse, maintain that in order to make the best possible use of available resources the techniques of production adopted should be 'labour intensive', that is, have a low capital/labour ratio. The choice of investment projects should therefore be made in such a way that productive resources are used in the same proportions in which they are available at the beginning of the process of industrialization, even if this will generally involve the adoption of a relatively backward technology.

This thesis has been contrasted with the following argument, put forward, among others, by Maurice Dobb: if a developing country adopted a capital intensive technology, then the immediate effect of investment on employment would not be as favourable as in the case of a labour intensive technology, because the amount of labour required per unit of capital would be less. But the larger stock of capital equipment per worker would cause labour to be more productive and if we assume, as seems reasonable, that the wage rate would be the same (that is, close to subsistence level) whichever the technology adopted, then the greater productivity would give rise to a greater surplus of production over consumption, and therefore to a larger flow of investment. This, in turn, would allow a faster rate of growth of income and therefore, in the long run, a higher level of employment than might have been allowed by the

adoption of a labour-intensive technology.

According to this argument under-developed countries would therefore be well advised to adopt, from the outset of the development process, advanced, capital-intensive techniques, at least in those sectors in which the productivity of labour is more sensitive to differences in technology. But the use of advanced technologies and the high rate of investment that these technologies require and make possible, render even more crucial the problem of the indivisibility of investments which, as we have seen, is characteristic of a situation of under-development.

All this means that, unlike what happens in already developed countries, in under-developed economies the flow of investment has to be very large in relation to the existing stock of capital, because of the problem of the indivisibility of investment. It follows that the risk of investing is much greater in under-developed than in already industrialized countries, and, therefore, that the initiation of a process of industrialization is generally not within the scope of private entrepreneurs. It is for this reason that most industrial developments in backward countries have been the result of government planning, with the government playing a decisive role, rather than of 'spontaneous' market forces. These development plans, however, are never as comprehensive and detailed as those adopted in communist countries, and their nature and content vary very widely from one country to another. In particular, the proportions in which private and public enterprise are combined vary considerably according to the local situation.

The tasks assigned to the government in a development plan are generally three. Firstly, the government is given over-all responsibility for the conduct of the development process. To do this it must first of all fix the over-all targets, generally expressed in terms of income and employment, that the country is supposed to reach within a given period of time. Then it must indicate the rate of investment necessary to reach those targets within the prescribed time. Finally, it is for the government to ensure that the flow of saving is sufficient to finance the required investment. This is essential if the vicious circle of poverty is to be broken, and is generally achieved by holding down the rate of growth of consumption so that an increasing fraction of income is saved and invested as income grows. A temporary sacrifice in consumption is, therefore, necessary in order to utilize the productive potential represented by the presence of concealed unemployment, which could not be mobilized without a much larger stock of capital and, consequently, a high level of investment and saving. The government's second

task is to eliminate all those institutional bottlenecks that stand in the way of the development process. This usually takes the form of land reform and of new arrangements (including nationalization) with foreign financial interests in the country. Thirdly, the government is expected to invest directly not only in infrastructures (roads, ports, etc.) but also in industrial capital, particularly in those basic industries whose existence is a pre-condition for the profitability of private investment.

The industrialization of a developing country is, therefore, characterized by elements of government planning and elements of private enterprise. The co-existence of these two elements is a source of tension and attrition, both of which have proved very difficult to eliminate. So far, economists have been unable to produce a theory of economic development based on private enterprise (and therefore on market mechanisms) the main lines of which have been fixed by the planning authorities, and in which the government operates also as an entrepreneur in order to create the preliminary conditions for industrial growth. The inadequacy of economic theory on this point is particularly disturbing, as the course that will be followed by today's under-developed countries is bound to change, perhaps soon, the very structure of world economy.

BIBLIOGRAPHICAL NOTE

The literature on developing countries and the economics of underdevelopment is vast and very specialized. A good textbook treatment is that of C. P. Kindleberger *Economic development* (McGraw-Hill, 1958, 2nd edition 1965); see also W. W. Rostow, *The stages of economic growth* (Cambridge University Press, 1960).

On the causes of international disparities see G. Myrdal's fundamental book, *Economic theory and underdeveloped regions* (Duckworth, 1957); also G. M. Meyer, *The international economics of development* (Harper & Row, 1968).

On the appropriate choice of techniques for a developing country see R. Nurkse, *Problems of capital formation in underdeveloped countries* (Oxford University Press, 1953); M. Dobb, *An essay on economic growth and planning* (Routledge, 1960), and 'On the theory of investment planning and economic growth' (in *Problems of economic dynamics and planning: essays in honour of M. Kalecki*, Pergamon, 1966); A. K. Sen, *Choice of techniques* (Blackwell, 3rd edition 1968).

Chapter 12

CONTEMPORARY PROBLEMS OF THE WORLD ECONOMY

1. *The situation in mature economies: the 'affluent society'*

The highly developed economies are those that, according to one thesis, have reached or are approaching a phase of maturity or stagnation. The functioning of these economies has changed substantially over the last hundred years. The classical picture of a capitalist economy, as seen by Smith and Marx, was one in which wages were systematically reduced to the bare subsistence level, so that any increase in productivity resulted in an increase in investment not only in absolute terms, but also as a proportion of national income. Capitalism, therefore, appeared as a form of organization having the sole aim of increasing production, and in which consumption was reduced to the role of an intermediate product, needed only in order to keep alive the labour force. This explains why the classical economists conceived profit not as the reward for a specific contribution to production, but as a residual, as what is left of income after the remuneration of labour. The function of such an economic system, then, was that of preserving and expanding the residual by compressing wages to subsistence level, and thus increasing the productivity of labour by increasing the stock of capital per worker.

There is no doubt that this was a correct interpretation of the economic conditions of the time, and that Marx was justified in pointing to the discrepancy between the productive capacity and purchasing power of the workers as the ultimate cause of periodical economic crises. This point had also been realized by Malthus, who, in contrast with Smith and Ricardo, maintained that the economic system could never reach a position of permanent equilibrium without a considerable expansion in what the classics called 'unproductive consumption'; that is, the consumption of the rich. Malthus saw, in fact, that the imbalance between production and consumption is intrinsic to the capitalist economy and, as he could

not conceive the consumption of the workers as being at more than subsistence level, he looked at the consumption of the rich as to a safety valve.

This classical picture of the capitalist economy has been gradually changing. First of all, the tendency to reduce wages to subsistence level has been brought to an end by the emergence and success of the workers' unions, as a result of which national income is no longer distributed according to the pure logic of capitalist expansion. The working class has thus escaped the process of progressive impoverishment predicted by Marx, which was only applicable to a society which could be called purely capitalistic—a society in which the worker, reduced to mere labour force, has no contractual power.

The distribution of income has been greatly affected also by the economic policies pursued by the governments of democratic countries. These policies consist of direct redistribution of income through taxes and subsidies, and of indirect redistribution resulting from anticyclical measures of the Keynesian type. All these factors have had considerable effects on the capitalistic reality. In the case of Britain, for example, Strachey has shown that the quota of earned incomes in the total remained constant up to 1939 and increased slightly after the Second World War.

As a result of these redistributions of income, the level of consumption in modern capitalist societies is now very high. But, although much larger, consumption has not been able to escape the logic that makes it an intermediate stage of production. Consumption is in fact no longer at subsistence level, but is still completely subordinated to production in the sense that it is not autonomously determined by the consumers, but by the producers. This explains, among other things, the crucial role played by advertising in modern economies.

From this particular aspect of consumption stem some of the most delicate problems of mature economies. Particularly in the United States, where the level of consumption is extremely high, the difficulty in creating new forms of consumption has now become evident. A very interesting analysis based on considerations of this kind is contained in Galbraith's book *The affluent society*. This author points out the irrationality of the capitalist economy even from the point of view of productive efficiency, which is generally regarded as its major strength.

According to the standard economic theory, the mechanism of profit maximization is sufficient to ensure the efficient satisfaction of consumer requirements. But the efficiency of this mechanism

could only be judged if its targets were really independent of the operation of the mechanism itself. It would be necessary to establish what consumer requirements are, independently of production, before we could judge whether the productive system satisfies them efficiently or not. But, once the elementary needs have been satisfied, the additional consumption is dictated by the producers, so this criterion of efficiency loses its meaning and one even starts to doubt the rationality of the economic process as a whole. Thus Galbraith regards as a myth the importance attributed by American society to the productive phenomenon. Moreover, this kind of irrational relationship between production and consumption does not contain any guarantee of sustained growth. On the contrary, phenomena of saturation are bound to become more and more frequent, as it becomes more and more difficult to manipulate the consumer in order to make a further expansion of production possible.

The problems posed by the affluent society are having political repercussions also, because their solution would require the expansion of the public sphere of influence at the expense of the private one. It is in fact suggested that a larger part of total demand should be under public control, so as to avoid the uncertainties and difficulties involved in expanding private consumption. As regards the American economy, increasing pressure is being put on the federal government to expand certain public services, for which there is not yet sufficient provision in spite of the general affluence enjoyed by the country. More public expenditure is in fact needed to provide the basic infrastructures still lacking in certain parts of the United States, and also to improve and expand the educational system. Along the same lines are the proposals that the United States and the other developed economies should devote a much larger fraction of their resources to helping the under-developed countries. The exact economic and historical significance of a policy of greater generosity towards the under-developed countries can be better understood by reconsidering the causes of international inequalities from one particular angle.

2. *Welfare and under-development*

The usual way of explaining the widening gap between developed and under-developed countries centres, as we have seen, on the risk attended to investing in areas that are not already industrialized.

The risk derives mainly from the fact that in an area not already industrialized an exceedingly large proportion of the total cost of an investment project consists of fixed costs, because the investment has to create its own external economies by establishing, all at once, a whole complex of interrelated industrial activities. The difficulty in extending the capitalistic mode of production to the underdeveloped world is therefore explained by a technological argument, analogous to the one used by Marx to justify his theory of the falling rate of profit. In practice, the risk of investing in underdeveloped countries has been considered bearable, and has actually been borne, only in connection with the exploitation of natural resources needed as inputs by the industries of the developed world.

There is, however, another aspect of the problem which is closely connected with what is going on in the developed world. The gradual expansion of capitalism to the whole world, predicted by the classics, was to come about as a result of its extraordinary expansionary thirst. This dynamism depended, in turn, on the fact that all increases in productivity resulted in further investment, because the wage rate was kept at subsistence level. If this state of affairs had persisted, it would have been necessary to find new geographical areas in which to invest the savings of the industrialized countries and, therefore, it is not unrealistic to think that capitalism would have expanded in this way to embrace the whole world.

If one accepts this argument, it becomes clear that the changes that have taken place in the organization of capitalist countries must have had considerable influence on the present imbalance in the world economic situation. The redistribution of income from the wealthy classes to the working classes that has taken place in mature economies, and the consequent changes in the division of income between consumption and saving, must have weakened the original expansionary force of the system. The capitalist world has become inward-looking, in the sense that more and more of its energies are directed towards increasing the welfare of its inhabitants. The chances of geographical expansion have become progressively smaller: part of the resources that in the old expansionist days would have gone to create new industrial developments in new areas, are now used for the consumption of the affluent society. There is, therefore, a close link between the 'welfare state', which has become the ideal of industrial societies, at least in Western Europe, and the backwardness of the countries that missed the original industrial revolution.

In the circumstances, it is natural that the idea of channelling

more of the resources of the richer countries towards the under-developed world should be reproposed whenever these resources remain unutilized because of a slowdown in the growth of consumer demand. But the problem of achieving a redistribution of income on a world scale can no longer be put in the same terms as a century ago. For all the reasons that we have seen, a transfer of resources to under-developed countries is no longer within the scope and profitability of the individual entrepreneur. The problem is now one of government policy, and many of the ideas put forward by Western economists and politicians and by international organizations show a decisive movement in this direction.

3. *The planned economies*

The problems of planned economies are not very dissimilar from those of mature capitalist economies. We have seen how the economists of the socialist countries are now resolutely orientated towards forms of economic planning, in which the planning authorities are expected to decide not only the broad lines of development of the system, but also its detailed productive configuration or the set of prices corresponding to it. The main economic problem is, therefore, in these countries too, one of efficient utilization of resources. The main difference between capitalist and socialist economies is that in the latter the emphasis is still on capital accumulation rather than on consumption. But, as the level of income *per capita* in the socialist economies tends to approach that of the richer capitalist countries, the targets of economic planning also tend to be modified in favour of consumption.

The recent experience of the Soviet Union is a case in point. Soviet 'economic ideals' are no longer what they were as recently as fifteen years ago: capital accumulation has been ousted from first place, and the official aim now is to attain the level of consumption *per capita* enjoyed by the richer Western countries. Assuming that the problem of the efficient management of a planned economy can be solved, we can therefore expect that the socialist economies too will eventually be faced with problems analogous to those of the affluent capitalist societies.

The reference to Marx is to Book III of *Das Kapital.* On Malthus, see Keynes' essay in *Essays in biography* (Macmillan, 1933).

On the situation in the mature economies see E. J. Strachey, *Contemporary capitalism* (Random House, 1956); J. K. Galbraith, *The affluent society* (Hamish Hamilton, 1958), *American capitalism* (Hamish Hamilton, 1963), *The new industrial state* (Hamish Hamilton, 1967). See also J. E. Meade 'Is "the new industrial state" inevitable?' (*Economic Journal*, 1968).

Chapter 13

RECENT TRENDS IN ECONOMIC RESEARCH

1. *'Classical' economics and 'modern' economics*

Having looked, however briefly, at the economic problems facing us today, we may ask what instruments of analysis economists now have at their disposal. This book as a whole is partly a reply to this question, but it is opportune, at this point, to attempt to bring out the essential characteristics of twentieth-century economics. This will be easier if we contrast the 'modern' approach with what is usually referred to as the 'classical' approach.

The difference between the two approaches is very great because it reflects two substantially different visions of the economic process. According to the classics, from Smith to Ricardo and Marx, the economic process is aimed at the production of a *surplus value*, that is, a surplus of the value of the product over the cost of production. This cost consists of the remuneration of labour and the reintegration of the other means of production utilized, which in turn embody the remuneration of the labour employed in their production. Therefore, the cost of production consists entirely of the remuneration of labour directly and indirectly employed, and the surplus value is that part of the value of the product which does not correspond to any *specific* contribution. The surplus derives from the fact that the labour employed in the productive process is *more than* sufficient to increase the value of the means of production by the value of the goods necessary for the subsistence of the workers; it therefore includes *surplus labour.*

The surplus value is the source of the rent paid to the landlord and the profit paid to the capitalist, and is also the source of finance for investment; that is, the condition for expansion of the productive system. This expansion will actually take place only if the surplus value is not used for what the classics called 'unproductive consumption' which, unlike the subsistence consumption of the workers employed in production, not only makes no contribution to the formation of a surplus, but diverts part of it from what was

considered the proper end of the economic process.

The only discordant voice on the matter was, as we have seen, that of Malthus, who considered unproductive consumption necessary to keep demand in step with productive capacity.

Because they interpreted reality in this way, the classical economists run into serious analytical difficulties. Without going into too many details which would take us outside the scope of this book, we must mention a few of these difficulties, for this will help us to understand some recent theoretical developments. Let us start from the concept of surplus, which the classics see as the object of economic activity. This surplus is not expressed in kind, as in a feudal society, but in money terms: in the capitalist economy the surplus is a surplus value. Therefore, it is not possible to explain the functioning of the economic system without a theory of value that explains how the market mechanism generates a surplus.

This task was performed by what has become known as the labour theory of value. The theory was first set out, in imperfect and ambiguous terms, by Smith, clearly stated by Ricardo, and finally formulated in great detail by Marx. Its central proposition is that the relative prices of commodities tend in the long run to be equal to the ratios between the quantities of labour directly and indirectly required to produce the commodities themselves, under normal productive conditions. The great strength of this suggestive theory lies in the fact that it aims at explaining simultaneously the phenomenon of surplus and the concept of market value which, taken together, explain the phenomenon of a surplus value. The surplus is accounted for by referring to the labour content of all commodities, which makes it possible to distinguish between necessary labour and surplus labour. The concept of value is accounted for by identifying the rate of exchange between two commodities with their relative labour content, which gives the rule for expressing the surplus as a surplus value.

The difficulty here arises from the fact that it cannot be demonstrated that relative prices coincide with the ratios between the labour contents of the commodities. This difficulty was seen by Marx in the following terms. If direct labour and indirect labour (or capital) are combined in different proportions in different commodities, as will generally be the case, then relative prices cannot be equal to the ratios between labour contents without violating the two fundamental equilibrium conditions of a market economy: equality between demand and supply of each commodity, and the same rate of profit in all economic activities.

Modern economists have got round this difficulty by repudiating

altogether the concept of surplus, and, therefore, of surplus value. Modern economic theory is in fact based on the view that behind each form of income there is a specific productive contribution, so that there cannot be any surplus. In Walras this is done by assuming that, at the beginning of the productive process, there is a collection of original factors of production, the productive services of which serve to justify the various forms of income. We have also seen what analytical difficulties one runs into when the quantities of productive factors available are assumed to be fixed: this is what has prevented the inclusion of a theory of capital in the general equilibrium scheme.

As a matter of fact, there has been no further attempt, since Walras' first formulation, to deal with the problem of accumulation in a general equilibrium framework. The Walrasian theory has been completed in many aspects, more often criticized (remember, in particular, the theories of imperfect competition and the Keynesian theory of employment), but its really fundamental difficulty has not even been broached. The only exception to this state of affairs is Von Neumann's model, which represents an alternative to the Walrasian approach because it portrays a perfectly circular productive process without original factors and final consumption. Von Neumann, therefore, eliminates the Walrasian difficulty by resorting to a type of analysis that can be traced back to the classical tradition.

2. *The 'production of commodities by means of commodities' according to Sraffa*

A conscious attempt to develop a method of economic analysis outside the 'modern' tradition is contained in Piero Sraffa's book *Production of commodities by means of commodities*, published in 1960. Sraffa's book is presented as a preface to a critique of the theory of marginal productivity; that is, of that theory which assumes the productivity of the services of given factors of production as the fundamental economic category. Sraffa develops his argument in three stages.

He starts by describing the functioning of a perfectly closed productive system, in which the same commodities appear as products and as means of production. Here the quantity produced of each good is exactly equal to the amount employed of it as a means of production, and all these quantities are assumed to be given. This

is therefore a subsistence economy, producing just what is necessary to reproduce itself in the same form and dimensions. The relative prices of the various commodities are simply set equal to the relative costs of production.

Sraffa then goes on to describe an economic system generating a surplus. The hypothesis that the same commodities are used as inputs and as outputs is maintained, but the technology is now such that the quantity produced of each good can be greater than or equal to the quantity of the same good used as a means of production. The value of the output can therefore exceed its cost, in which case we have a surplus. This scheme is expressed in the form of a set of equations, which determine simultaneously the relative prices and the general profit rate. The presence of a rate of profit is a manifestation of the fact that the surplus (or profit) is distributed among the various productive activities in proportion to the value of the means of production employed in each of them. Here labour does not appear explicitly, but only indirectly in the form of goods consumed by the workers—goods that are regarded as means of production like all the others.

Finally, Sraffa undertakes to isolate labour from the other means of production and, therefore, to distinguish the wage rate from the other prices. This could be done in two different ways. In order to remain fully within the logic of his initial approach, the author should—as he remarks himself—distinguish between two components of the wage: one just sufficient for the subsistence of the worker, which should be subtracted from the value of production in order to obtain the surplus, and one representing the share of the worker in the division of the surplus produced by the system. But Sraffa prefers to adhere to the modern convention of regarding the wage as one thing. The net product, consisting of the amounts of goods produced minus those re-employed as means of production, therefore divides between wages and profits. But the profit rate and wage rate cannot both be determined by the equations of the system; one of these two magnitudes has to be determined from outside. Here Sraffa chooses to follow the classical tradition and assumes that the wage rate is determined by factors outside the economic system (on the basis of the historically determined level of subsistence and of the ability of workers to share in the appropriation of the surplus), so that the rate of profit becomes a function of the wage rate.

A peculiarity of Sraffa's approach is the very special unit that he uses to measure values. In order to understand the need for this particular unit of measurement, we must refer back to the attempt

made by Ricardo to explain the rate of profit of the economy as a whole by the rate of profit in agriculture: if wheat were the only agricultural product and the wages of agricultural workers consisted of wheat, then the agricultural rate of profit could be calculated in physical terms as the ratio between the agricultural surplus and the total wage fund anticipated to the workers, without any need to bring in prices, for both numerator and denominator would consist of wheat. In this case the agricultural rate of profit would not vary with relative prices, but only with changes in the real wage rate. But, as the rate of profit must be the same for all productive activities, it follows that in equilibrium the price system must be such that the profit rates of the various sectors of the economy are equal to the agricultural one.

In practice, however, output and capital (the wage fund) are not homogeneous even in agriculture, and the profit rate cannot be calculated without reference to prices. The labour theory of value offers an alternative way, because it allows the rate of profit to be calculated as the ratio between the quantity of labour contained in the goods forming the surplus and the quantity of labour contained in wage goods. There is no doubt that one of the purposes of the labour theory of value was indeed that of measuring the profit rate directly, without resorting to market prices. But, as the labour theory of value does not stand, this alternative way out is also closed.

Let us now look at Sraffa's approach. He constructs a composite commodity comprising individual goods combined in the same proportions in which they must be combined to produce the composite commodity itself. The advantage of defining this artificial commodity, which Sraffa calls the *standard commodity*, is that the product and the means of production are physically homogeneous. Consequently, the rate of profit can be calculated in physical terms for the standard commodity, as for Ricardo's wheat.

Having defined the standard commodity, Sraffa chooses as his unit of measurement for wages and prices the amount of standard commodity that could be produced by employing all the labour available in the system over a unit of time: this quantity is therefore set equal to one. Ricardo's problem is then completely solved: the rate of profit can be calculated in physical terms (that is, in units of standard commodity) for the standard commodity and, in equilibrium, the price system will have to be such that the rate of profit in the production of all individual commodities is equal to the one of the standard commodity.

Sraffa's analysis therefore departs substantially from the

'modern' line of thought and constitutes an explicit return to the classical tradition of Ricardo and Marx. His concept of profit is not that of the remuneration of a particular productive factor, but of a residual. His concept of wage rate has no connection whatever with the marginal productivity of labour, and is determined by social and historical mechanisms. Even consumption, far from being regarded as the aim of production as in the modern sense, is treated as a means towards the realization of a surplus.

How did Sraffa manage to avoid the contradictions that are in Ricardo and Marx, and still remain within the classical tradition?

The immediate answer is that Sraffa completely avoids the labour theory of value which, as is more evident in Marx, is at the root of the formal difficulties of classical economics. How different this makes Sraffa's theory from the classical one is one of the most interesting questions raised by Sraffa's work. The following considerations should be taken as simple indications of possible lines of research on the subject. The classical theory of value, although not completely coherent, had the advantage of explaining the surplus with reference to a clearly recognizable social phenomenon, which Marx explicitly called *exploitation.* The existence of a situation in which the available technology allows more to be produced than is necessary for the survival and reproduction of the workers employed is taken as a historical fact, which cannot be investigated any further. To this it could be objected that there could not be exploitation if the workers were not able, in a technical sense, to produce more than is necessary for their subsistence, and that, therefore, a certain rate of exploitation is the effect of a certain level of productiivty. But, as least for Marx, it is clear that this particular level of productivity could be reached only in a society divided into exploiters and exploited. Whether this idea of exploitation is correct or not is another matter, but what we want to stress here is how classical economics took as a starting-point a certain social reality, and how the labour theory of value provides the link between the economic system and this human reality. Because Sraffa, on the other hand, takes the material world of commodities and technology as his starting-point, he cannot explain it any further, as the classics attempted to do.

Moreover, Marx saw trade not only as an exchange of commodities but first and foremost as the establishment of a relationship between people, as the way in which people form society, or at least one particular type of society. The labour theory of value, by reducing the rate of exchange between commodities to their relative labour content, is there to remind us that the market con-

sists of people. Marx's prices are, therefore, without any doubt, market prices, the result of an act of exchange, while Sraffa's prices could very well be accounting prices. But, if the existence of individuals competing with one another in a market is not essential to the theory, how can the existence of a general rate of profit be justified? In Von Neumann the rate of profit can be justified on efficiency grounds, but the same is not true in Sraffa. Analytical rigour is, therefore, gained at the expense of social significance. Nevertheless, the possibility of rigorously describing a system aimed at the production of a surplus may prove useful to the understanding of the crucial relationship between production and consumption in modern industrial societies.

BIBLIOGRAPHICAL NOTE

Marx's theory of value is in Book III of *Das Kapital*; see also J. Robinson, *An essay on Marxian economics* (Macmillan, 1942, 2nd edition 1947). The book by P. Sraffa referred to in the text is *Production of commodities by means of commodities* (Cambridge University Press, 1960).

A fundamental contribution to the current discussion on the relevance of modern economic theory to the understanding and controlling of modern economic reality is a book by A. Lowe, *On economic knowledge: toward a science of political economics* (Harper & Row, 1965); see also R. L. Heilbroner (editor), *Economic means and social ends* (Prentice-Hall, 1969), containing papers presented by philosophers and social scientists to the two Adolph Lowe Symposia held in 1968 at the New School for Social Research in New York.