THE COLLECTED WRITINGS OF

JOHN MAYNARD KEYNES

VOLUME VII

THE GENERAL THEORY OF EMPLOYMENT INTEREST AND MONEY

MACMILLAN
CAMBRIDGE UNIVERSITY PRESS
FOR THE
ROYAL ECONOMIC SOCIETY

Chapter 1

THE GENERAL THEORY

I have called this book the *General Theory of Employment, Interest and Money*, placing the emphasis on the prefix *general*. The object of such a title is to contrast the character of my arguments and conclusions with those of the *classical*[1] theory of the subject, upon which I was brought up and which dominates the economic thought, both practical and theoretical, of the governing and academic classes of this generation, as it has for a hundred years past. I shall argue that the postulates of the classical theory are applicable to a special case only and not to the general case, the situation which it assumes being a limiting point of the possible positions of equilibrium. Moreover, the characteristics of the special case assumed by the classical theory happen not to be those of the economic society in which we actually live, with the result that its teaching is misleading and disastrous if we attempt to apply it to the facts of experience.

[1] 'The classical economists' was a name invented by Marx to cover Ricardo and James Mill and their *predecessors*, that is to say for the founders of the theory which culminated in the Ricardian economics. I have become accustomed, perhaps perpetrating a solecism, to include in 'the classical school' the *followers* of Ricardo, those, that is to say, who adopted and perfected the theory of the Ricardian economics, including (for example) J. S. Mill, Marshall, Edgeworth and Prof. Pigou.

Chapter 2

THE POSTULATES OF THE CLASSICAL ECONOMICS

Most treatises on the theory of value and production are primarily concerned with the distribution of a *given* volume of employed resources between different uses and with the conditions which, assuming the employment of this quantity of resources, determine their relative rewards and the relative values of their products.[1]

The question, also, of the volume of the *available* resources, in the sense of the size of the employable population, the extent of natural wealth and the accumulated capital equipment, has often been treated descriptively. But the pure theory of what determines the *actual employment* of the available resources has seldom been examined in great detail. To say that it has not been examined at all would, of course, be absurd. For every discussion concerning fluctuations of employment, of which there have been many, has been concerned with it. I mean, not that the topic has been overlooked, but that the fundamental theory

[1] This is in the Ricardian tradition. For Ricardo expressly repudiated any interest in the *amount* of the national dividend, as distinct from its distribution. In this he was assessing correctly the character of his own theory. But his successors, less clear-sighted, have used the classical theory in discussions concerning the causes of wealth. *Vide* Ricardo's letter to Malthus of October 9, 1820: 'Political Economy you think is an enquiry into the nature and causes of wealth—I think it should be called an enquiry into the laws which determine the division of the produce of industry amongst the classes who concur in its formation. No law can be laid down respecting quantity, but a tolerably correct one can be laid down respecting proportions. Every day I am more satisfied that the former enquiry is vain and delusive, and the latter only the true objects of the science.'

underlying it has been deemed so simple and obvious that it has received, at the most, a bare mention.[1]

I

The classical theory of employment—supposedly simple and obvious—has been based, I think, on two fundamental postulates, though practically without discussion, namely:

I. *The wage is equal to the marginal product of labour*

That is to say, the wage of an employed person is equal to the value which would be lost if employment were to be reduced by one unit (after deducting any other costs which this reduction of output would avoid); subject, however, to the qualification that the equality may be disturbed, in accordance with certain principles, if competition and markets are imperfect.

II. *The utility of the wage when a given volume of labour is employed is equal to the marginal disutility of that amount of employment.*

That is to say, the real wage of an employed person is that which is just sufficient (in the estimation of the employed persons themselves) to induce the volume of labour actually employed to be forthcoming; subject to the qualification that the equality for each individual unit of labour may be disturbed by combination between employable units analogous to the imperfections

[1] For example, Prof. Pigou in the *Economics of Welfare* (4th ed. p. 127) writes (my italics): 'Throughout this discussion, except when the contrary is expressly stated, the fact that some resources are generally unemployed against the will of the owners is ignored. *This does not affect the substance of the argument*, while it simplifies its exposition.' Thus, whilst Ricardo expressly disclaimed any attempt to deal with the amount of the national dividend as a whole, Prof. Pigou, in a book which is specifically directed to the problem of the national dividend, maintains that the same theory holds good when there is some involuntary unemployment as in the case of full employment.

of competition which qualify the first postulate. Disutility must be here understood to cover every kind of reason which might lead a man, or a body of men, to withhold their labour rather than accept a wage which had to them a utility below a certain minimum.

This postulate is compatible with what may be called 'frictional' unemployment. For a realistic interpretation of it legitimately allows for various inexactnesses of adjustment which stand in the way of continuous full employment: for example, unemployment due to a temporary want of balance between the relative quantities of specialised resources as a result of miscalculation or intermittent demand; or to time-lags consequent on unforeseen changes; or to the fact that the change-over from one employment to another cannot be effected without a certain delay, so that there will always exist in a non-static society a proportion of resources unemployed 'between jobs'. In addition to 'frictional' unemployment, the postulate is also compatible with 'voluntary' unemployment due to the refusal or inability of a unit of labour, as a result of legislation or social practices or of combination for collective bargaining or of slow response to change or of mere human obstinacy, to accept a reward corresponding to the value of the product attributable to its marginal productivity. But these two categories of 'frictional' unemployment and 'voluntary' unemployment are comprehensive. The classical postulates do not admit of the possibility of the third category, which I shall define below as 'involuntary' unemployment.

Subject to these qualifications, the volume of employed resources is duly determined, according to the classical theory, by the two postulates. The first gives us the demand schedule for employment, the second gives us the supply schedule; and the amount of employment is fixed at the point where the utility of the marginal product balances the disutility of the marginal employment.

It would follow from this that there are only four possible means of increasing employment:

(*a*) An improvement in organisation or in foresight which diminishes 'frictional' unemployment;

(*b*) a decrease in the marginal disutility of labour, as expressed by the real wage for which additional labour is available, so as to diminish 'voluntary' unemployment;

(*c*) an increase in the marginal physical productivity of labour in the wage-goods industries (to use Professor Pigou's convenient term for goods upon the price of which the utility of the money-wage depends);

or (*d*) an increase in the price of non-wage-goods compared with the price of wage-goods, associated with a shift in the expenditure of non-wage-earners from wage-goods to non-wage-goods.

This, to the best of my understanding, is the substance of Professor Pigou's *Theory of Unemployment* —the only detailed account of the classical theory of employment which exists.[1]

II

Is it true that the above categories are comprehensive in view of the fact that the population generally is seldom doing as much work as it would like to do on the basis of the current wage? For, admittedly, more labour would, as a rule, be forthcoming at the existing money-wage if it were demanded.[2] The classical school reconcile this phenomenon with their second postulate by arguing that, while the demand for labour

[1] Prof. Pigou's *Theory of Unemployment* is examined in more detail in the Appendix to Chapter 19 below.

[2] Cf. the quotation from Prof. Pigou above, p. 5, footnote.

at the existing money-wage may be satisfied before everyone willing to work at this wage is employed, this situation is due to an open or tacit agreement amongst workers not to work for less, and that if labour as a whole would agree to a reduction of money-wages more employment would be forthcoming. If this is the case, such unemployment, though apparently involuntary, is not strictly so, and ought to be included under the above category of 'voluntary' unemployment due to the effects of collective bargaining, etc.

This calls for two observations, the first of which relates to the actual attitude of workers towards real wages and money-wages respectively and is not theoretically fundamental, but the second of which is fundamental.

Let us assume, for the moment, that labour is not prepared to work for a lower money-wage and that a reduction in the existing level of money-wages would lead, through strikes or otherwise, to a withdrawal from the labour market of labour which is now employed. Does it follow from this that the existing level of real wages accurately measures the marginal disutility of labour? Not necessarily. For, although a reduction in the existing money-wage would lead to a withdrawal of labour, it does not follow that a fall in the value of the existing money-wage in terms of wage-goods would do so, if it were due to a rise in the price of the latter. In other words, it may be the case that within a certain range the demand of labour is for a minimum money-wage and not for a minimum real wage. The classical school have tacitly assumed that this would involve no significant change in their theory. But this is not so. For if the supply of labour is not a function of real wages as its sole variable, their argument breaks down entirely and leaves the question of what the actual employment will be quite indeterminate.[1] They do not seem to have realised that, unless the supply of labour is a function of real wages alone,

[1] This point is dealt with in detail in the Appendix to chapter 19 below.

their supply curve for labour will shift bodily with every movement of prices. Thus their method is tied up with their very special assumptions, and cannot be adapted to deal with the more general case.

Now ordinary experience tells us, beyond doubt, that a situation where labour stipulates (within limits) for a money-wage rather than a real wage, so far from being a mere possibility, is the normal case. Whilst workers will usually resist a reduction of money-wages, it is not their practice to withdraw their labour whenever there is a rise in the price of wage-goods. It is sometimes said that it would be illogical for labour to resist a reduction of money-wages but not to resist a reduction of real wages. For reasons given below (p. 14), this might not be so illogical as it appears at first; and, as we shall see later, fortunately so. But, whether logical or illogical, experience shows that this is how labour in fact behaves.

Moreover, the contention that the unemployment which characterises a depression is due to a refusal by labour to accept a reduction of money-wages is not clearly supported by the facts. It is not very plausible to assert that unemployment in the United States in 1932 was due either to labour obstinately refusing to accept a reduction of money-wages or to its obstinately demanding a real wage beyond what the productivity of the economic machine was capable of furnishing. Wide variations are experienced in the volume of employment without any apparent change either in the minimum real demands of labour or in its productivity. Labour is not more truculent in the depression than in the boom—far from it. Nor is its physical productivity less. These facts from experience are a *prima facie* ground for questioning the adequacy of the classical analysis.

It[1] would be interesting to see the results of a statistical enquiry into the actual relationship between

[1] See Appendix 3 below for further discussion on this point. [Ed.]

changes in money-wages and changes in real wages. In the case of a change peculiar to a particular industry one would expect the change in real wages to be in the same direction as the change in money-wages. But in the case of changes in the general level of wages, it will be found, I think, that the change in real wages associated with a change in money-wages, so far from being usually in the same direction, is almost always in the opposite direction. When money-wages are rising, that is to say, it will be found that real wages are falling; and when money-wages are falling, real wages are rising. This is because, in the short period, falling money-wages and rising real wages are each, for independent reasons, likely to accompany decreasing employment; labour being readier to accept wage-cuts when employment is falling off, yet real wages inevitably rising in the same circumstances on account of the increasing marginal return to a given capital equipment when output is diminished.

If, indeed, it were true that the existing real wage is a minimum below which more labour than is now employed will not be forthcoming in any circumstances, involuntary unemployment, apart from frictional unemployment, would be non-existent. But to suppose that this is invariably the case would be absurd. For more labour than is at present employed is usually available at the existing money-wage, even though the price of wage-goods is rising and, consequently, the real wage falling. If this is true, the wage-goods equivalent of the existing money-wage is not an accurate indication of the marginal disutility of labour, and the second postulate does not hold good.

But there is a more fundamental objection. The second postulate flows from the idea that the real wages of labour depend on the wage bargains which labour makes with the entrepreneurs. It is admitted, of course, that the bargains are actually made in terms of money, and even that the real wages acceptable to labour are

not altogether independent of what the corresponding money-wage happens to be. Nevertheless it is the money-wage thus arrived at which is held to determine the real wage. Thus the classical theory assumes that it is always open to labour to reduce its real wage by accepting a reduction in its money-wage. The postulate that there is a tendency for the real wage to come to equality with the marginal disutility of labour clearly presumes that labour itself is in a position to decide the real wage for which it works, though not the quantity of employment forthcoming at this wage.

The traditional theory maintains, in short, *that the wage bargains between the entrepreneurs and the workers determine the real wage*; so that, assuming free competition amongst employers and no restrictive combination amongst workers, the latter can, if they wish, bring their real wages into conformity with the marginal disutility of the amount of employment offered by the employers at that wage. If this is not true, then there is no longer any reason to expect a tendency towards equality between the real wage and the marginal disutility of labour.

The classical conclusions are intended, it must be remembered, to apply to the whole body of labour and do not mean merely that a single individual can get employment by accepting a cut in money-wages which his fellows refuse. They are supposed to be equally applicable to a closed system as to an open system, and are not dependent on the characteristics of an open system or on the effects of a reduction of money-wages in a single country on its foreign trade, which lie, of course, entirely outside the field of this discussion. Nor are they based on indirect effects due to a lower wages-bill in terms of money having certain reactions on the banking system and the state of credit, effects which we shall examine in detail in chapter 19. They are based on the belief that in a closed system a reduction

in the general level of money-wages will be accompanied, at any rate in the short period and subject only to minor qualifications, by some, though not always a proportionate, reduction in real wages.

Now the assumption that the general level of real wages depends on the money-wage bargains between the employers and the workers is not obviously true. Indeed it is strange that so little attempt should have been made to prove or to refute it. For it is far from being consistent with the general tenor of the classical theory, which has taught us to believe that prices are governed by marginal prime cost in terms of money and that money-wages largely govern marginal prime cost. Thus if money-wages change, one would have expected the classical school to argue that prices would change in almost the same proportion, leaving the real wage and the level of unemployment practically the same as before, any small gain or loss to labour being at the expense or profit of other elements of marginal cost which have been left unaltered.[1] They seem, however, to have been diverted from this line of thought, partly by the settled conviction that labour is in a position to determine its own real wage and partly, perhaps, by preoccupation with the idea that prices depend on the quantity of money. And the belief in the proposition that labour is always in a position to determine its own real wage, once adopted, has been maintained by its being confused with the proposition that labour is always in a position to determine what real wage shall correspond to *full* employment, i.e. the *maximum* quantity of employment which is compatible with a given real wage.

To sum up: there are two objections to the second postulate of the classical theory. The first relates to the actual behaviour of labour. A fall in real wages due

[1] This argument would, indeed, contain, to my thinking, a large element of truth, though the complete results of a change in money-wages are more complex, as we shall show in chapter 19 below.

to a rise in prices, with money-wages unaltered, does not, as a rule, cause the supply of available labour on offer at the current wage to fall below the amount actually employed prior to the rise of prices. To suppose that it does is to suppose that all those who are now unemployed though willing to work at the current wage will withdraw the offer of their labour in the event of even a small rise in the cost of living. Yet this strange supposition apparently underlies Professor Pigou's *Theory of Unemployment*,[1] and it is what all members of the orthodox school are tacitly assuming.

But the other, more fundamental, objection, which we shall develop in the ensuing chapters, flows from our disputing the assumption that the general level of real wages is directly determined by the character of the wage bargain. In assuming that the wage bargain determines the real wage the classical school have slipt in an illicit assumption. For there may be *no* method available to labour as a whole whereby it can bring the wage-goods equivalent of the general level of money-wages into conformity with the marginal disutility of the current volume of employment. There may exist no expedient by which labour as a whole can reduce its *real* wage to a given figure by making revised *money* bargains with the entrepreneurs. This will be our contention. We shall endeavour to show that primarily it is certain other forces which determine the general level of real wages. The attempt to elucidate this problem will be one of our main themes. We shall argue that there has been a fundamental misunderstanding of how in this respect the economy in which we live actually works.

III

Though the struggle over money-wages between individuals and groups is often believed to determine

[1] Cf. chapter 19, Appendix.

the general level of real wages, it is, in fact, concerned with a different object. Since there is imperfect mobility of labour, and wages do not tend to an exact equality of net advantage in different occupations, any individual or group of individuals, who consent to a reduction of money-wages relatively to others, will suffer a *relative* reduction in real wages, which is a sufficient justification for them to resist it. On the other hand it would be impracticable to resist every reduction of real wages, due to a change in the purchasing-power of money which affects all workers alike; and in fact reductions of real wages arising in this way are not, as a rule, resisted unless they proceed to an extreme degree. Moreover, a resistance to reductions in money-wages applying to particular industries does not raise the same insuperable bar to an increase in aggregate employment which would result from a similar resistance to every reduction in real wages.

In other words, the struggle about money-wages primarily affects the *distribution* of the aggregate real wage between different labour-groups, and not its average amount per unit of employment, which depends, as we shall see, on a different set of forces. The effect of combination on the part of a group of workers is to protect their *relative* real wage. The *general* level of real wages depends on the other forces of the economic system.

Thus it is fortunate that the workers, though unconsciously, are instinctively more reasonable economists than the classical school, inasmuch as they resist reductions of money-wages, which are seldom or never of an all-round character, even though the existing real equivalent of these wages exceeds the marginal disutility of the existing employment; whereas they do not resist reductions of real wages, which are associated with increases in aggregate employment and leave relative money-wages unchanged, unless the reduction proceeds so far as to threaten a reduction of the real

wage below the marginal disutility of the existing volume of employment. Every trade union will put up some resistance to a cut in money-wages, however small. But since no trade union would dream of striking on every occasion of a rise in the cost of living, they do not raise the obstacle to any increase in aggregate employment which is attributed to them by the classical school.

IV

We must now define the third category of unemployment, namely 'involuntary' unemployment in the strict sense, the possibility of which the classical theory does not admit.

Clearly we do not mean by 'involuntary' unemployment the mere existence of an unexhausted capacity to work. An eight-hour day does not constitute unemployment because it is not beyond human capacity to work ten hours. Nor should we regard as 'involuntary' unemployment the withdrawal of their labour by a body of workers because they do not choose to work for less than a certain real reward. Furthermore, it will be convenient to exclude 'frictional' unemployment from our definition of 'involuntary' unemployment. My definition is, therefore, as follows: *Men are involuntarily unemployed if, in the event of a small rise in the price of wage-goods relatively to the money-wage, both the aggregate supply of labour willing to work for the current money-wage and the aggregate demand for it at that wage would be greater than the existing volume of employment.* An alternative definition, which amounts, however, to the same thing, will be given in the next chapter (p. 26 below).

It follows from this definition that the equality of the real wage to the marginal disutility of employment presupposed by the second postulate, realistically interpreted, corresponds to the absence of 'involuntary' unemployment. This state of affairs we shall describe

as 'full' employment, both 'frictional' and 'voluntary' unemployment being consistent with 'full' employment thus defined. This fits in, we shall find, with other characteristics of the classical theory, which is best regarded as a theory of distribution in conditions of full employment. So long as the classical postulates hold good, unemployment, which is in the above sense involuntary, cannot occur. Apparent unemployment must, therefore, be the result either of temporary loss of work of the 'between jobs' type or of intermittent demand for highly specialised resources or of the effect of a trade union 'closed shop' on the employment of free labour. Thus writers in the classical tradition, overlooking the special assumption underlying their theory, have been driven inevitably to the conclusion, perfectly logical on their assumption, that apparent unemployment (apart from the admitted exceptions) must be due at bottom to a refusal by the unemployed factors to accept a reward which corresponds to their marginal productivity. A classical economist may sympathise with labour in refusing to accept a cut in its money-wage, and he will admit that it may not be wise to make it to meet conditions which are temporary; but scientific integrity forces him to declare that this refusal is, nevertheless, at the bottom of the trouble.

Obviously, however, if the classical theory is only applicable to the case of full employment, it is fallacious to apply it to the problems of involuntary unemployment—if there be such a thing (and who will deny it?). The classical theorists resemble Euclidean geometers in a non-Euclidean world who, discovering that in experience straight lines apparently parallel often meet, rebuke the lines for not keeping straight—as the only remedy for the unfortunate collisions which are occurring. Yet, in truth, there is no remedy except to throw over the axiom of parallels and to work out a non-Euclidean geometry. Something similar is required to-day in economics. We need to throw over

the second postulate of the classical doctrine and to work out the behaviour of a system in which involuntary unemployment in the strict sense is possible.

V

In emphasising our point of departure from the classical system, we must not overlook an important point of agreement. For we shall maintain the first postulate as heretofore, subject only to the same qualifications as in the classical theory; and we must pause, for a moment, to consider what this involves.

It means that, with a given organisation, equipment and technique, real wages and the volume of output (and hence of employment) are uniquely correlated, so that, in general, an increase in employment can only occur to the accompaniment of a decline in the rate of real wages. Thus I am not disputing this vital fact which the classical economists have (rightly) asserted as indefeasible. In a given state of organisation, equipment and technique, the real wage earned by a unit of labour has a unique (inverse) correlation with the volume of employment. Thus *if* employment increases, then, in the short period, the reward per unit of labour in terms of wage-goods must, in general, decline and profits increase.[1] This is simply the obverse of the familiar proposition that industry is normally working subject to decreasing returns in the short period during which equipment etc. is assumed to be constant; so that the marginal product in the wage-good industries (which governs real wages) neces-

[1] The argument runs as follows: n men are employed, the nth man adds a bushel a day to the harvest, and wages have a buying power of a bushel a day. The $n+1$th man, however, would only add ·9 bushel a day, and employment cannot, therefore, rise to $n+1$ men unless the price of corn rises relatively to wages until daily wages have a buying power of ·9 bushel. Aggregate wages would then amount to $\frac{9}{10}(n+1)$ bushels as compared with n bushels previously. Thus the employment of an additional man will, if it occurs, necessarily involve a transfer of income from those previously in work to the entrepreneurs.

that the aggregate demand price is equal to the aggregate supply price for all levels of output and employment.

These three assumptions, however, all amount to the same thing in the sense that they all stand and fall together, any one of them logically involving the other two.

Chapter 3

THE PRINCIPLE OF EFFECTIVE DEMAND

I

We need, to start with, a few terms which will be defined precisely later. In a given state of technique, resources and costs, the employment of a given volume of labour by an entrepreneur involves him in two kinds of expense: first of all, the amounts which he pays out to the factors of production (exclusive of other entrepreneurs) for their current services, which we shall call the *factor cost* of the employment in question; and secondly, the amounts which he pays out to other entrepreneurs for what he has to purchase from them together with the sacrifice which he incurs by employing the equipment instead of leaving it idle, which we shall call the *user cost* of the employment in question.[1] The excess of the value of the resulting output over the sum of its factor cost and its user cost is the profit or, as we shall call it, the *income* of the entrepreneur. The factor cost is, of course, the same thing, looked at from the point of view of the entrepreneur, as what the factors of production regard as their income. Thus the factor cost and the entrepreneur's profit make up, between them, what we shall define as the *total income* resulting from the employment given by the entrepreneur. The entrepreneur's profit thus defined is, as it should be, the quantity which he endeavours to maximise when he is deciding what amount of employ-

[1] A precise definition of *user cost* will be given in chapter 6.

ment to offer. It is sometimes convenient, when we are looking at it from the entrepreneur's standpoint, to call the aggregate income (i.e. factor cost *plus* profit) resulting from a given amount of employment the *proceeds* of that employment. On the other hand, the aggregate supply price[1] of the output of a given amount of employment is the expectation of proceeds which will just make it worth the while of the entrepreneurs to give that employment.[2]

It follows that in a given situation of technique, resources and factor cost per unit of employment, the amount of employment, both in each individual firm and industry and in the aggregate, depends on the amount of the proceeds which the entrepreneurs expect to receive from the corresponding output.[3] For entrepreneurs will endeavour to fix the amount of employ-

[1] Not to be confused (*vide infra*) with the supply price of a unit of output in the ordinary sense of this term.

[2] The reader will observe that I am deducting the user cost both from the *proceeds* and from the *aggregate supply price* of a given volume of output, so that both these terms are to be interpreted *net* of user cost; whereas the aggregate sums paid by the purchasers are, of course, *gross* of user cost. The reasons why this is convenient will be given in chapter 6. The essential point is that the aggregate proceeds and aggregate supply price net of user cost can be defined uniquely and unambiguously; whereas, since user cost is obviously dependent both on the degree of integration of industry and on the extent to which entrepreneurs buy from one another, there can be no definition of the aggregate sums paid by purchasers, *inclusive* of user cost, which is independent of these factors. There is a similar difficulty even in defining supply price in the ordinary sense for an individual producer; and in the case of the aggregate supply price of *output as a whole* serious difficulties of duplication are involved, which have not always been faced. If the term is to be interpreted gross of user cost, they can only be overcome by making special assumptions relating to the integration of entrepreneurs in groups according as they produce consumption-goods or capital-goods, which are obscure and complicated in themselves and do not correspond to the facts. If, however, aggregate supply price is defined as above *net* of user cost, these difficulties do not arise. The reader is advised, however, to await the fuller discussion in chapter 6 and its Appendix.

[3] An entrepreneur, who has to reach a practical decision as to his scale of production, does not, of course, entertain a single undoubting expectation of what the sale-proceeds of a given output will be, but several hypothetical expectations held with varying degrees of probability and definiteness. By his expectation of proceeds I mean, therefore, that expectation of proceeds which, if it were held with certainty, would lead to the same behaviour as does the bundle of vague and more various possibilities which actually makes up his state of expectation when he reaches his decision.

ment at the level which they expect to maximise the excess of the proceeds over the factor cost.

Let Z be the aggregate supply price of the output from employing N men, the relationship between Z and N being written $Z = \phi(N)$, which can be called the *aggregate supply function.*[1] Similarly, let D be the proceeds which entrepreneurs expect to receive from the employment of N men, the relationship between D and N being written $D = f(N)$, which can be called the *aggregate demand function.*

Now if for a given value of N the expected proceeds are greater than the aggregate supply price, i.e. if D is greater than Z, there will be an incentive to entrepreneurs to increase employment beyond N and, if necessary, to raise costs by competing with one another for the factors of production, up to the value of N for which Z has become equal to D. Thus the volume of employment is given by the point of intersection between the aggregate demand function and the aggregate supply function; for it is at this point that the entrepreneurs' expectation of profits will be maximised. The value of D at the point of the aggregate demand function, where it is intersected by the aggregate supply function, will be called *the effective demand.* Since this is the substance of the General Theory of Employment, which it will be our object to expound, the succeeding chapters will be largely occupied with examining the various factors upon which these two functions depend.

The classical doctrine, on the other hand, which used to be expressed categorically in the statement that 'Supply creates its own Demand' and continues to underlie all orthodox economic theory, involves a special assumption as to the relationship between these two functions. For 'Supply creates its own Demand' must mean that $f(N)$ and $\phi(N)$ are equal for *all* values

[1] In chapter 20 a function closely related to the above will be called the employment function.

of N, i.e. for all levels of output and employment; and that when there is an increase in $Z(=\phi(N))$ corresponding to an increase in N, $D(=f(N))$ necessarily increases by the same amount as Z. The classical theory assumes, in other words, that the aggregate demand price (or proceeds) always accommodates itself to the aggregate supply price; so that, whatever the value of N may be, the proceeds D assume a value equal to the aggregate supply price Z which corresponds to N. That is to say, effective demand, instead of having a unique equilibrium value, is an infinite range of values all equally admissible; and the amount of employment is indeterminate except in so far as the marginal disutility of labour sets an upper limit.

If this were true, competition between entrepreneurs would always lead to an expansion of employment up to the point at which the supply of output as a whole ceases to be elastic, i.e. where a further increase in the value of the effective demand will no longer be accompanied by any increase in output. Evidently this amounts to the same thing as full employment. In the previous chapter we have given a definition of full employment in terms of the behaviour of labour. An alternative, though equivalent, criterion is that at which we have now arrived, namely a situation in which aggregate employment is inelastic in response to an increase in the effective demand for its output. Thus Say's law, that the aggregate demand price of output as a whole is equal to its aggregate supply price for all volumes of output, is equivalent to the proposition that there is no obstacle to full employment. If, however, this is not the true law relating the aggregate demand and supply functions, there is a vitally important chapter of economic theory which remains to be written and without which all discussions concerning the volume of aggregate employment are futile.

II

A brief summary of the theory of employment to be worked out in the course of the following chapters may, perhaps, help the reader at this stage, even though it may not be fully intelligible. The terms involved will be more carefully defined in due course. In this summary we shall assume that the money-wage and other factor costs are constant per unit of labour employed. But this simplification, with which we shall dispense later, is introduced solely to facilitate the exposition. The essential character of the argument is precisely the same whether or not money-wages, etc., are liable to change.

The outline of our theory can be expressed as follows. When employment increases, aggregate real income is increased. The psychology of the community is such that when aggregate real income is increased aggregate consumption is increased, but not by so much as income. Hence employers would make a loss if the whole of the increased employment were to be devoted to satisfying the increased demand for immediate consumption. Thus, to justify any given amount of employment there must be an amount of current investment sufficient to absorb the excess of total output over what the community chooses to consume when employment is at the given level. For unless there is this amount of investment, the receipts of the entrepreneurs will be less than is required to induce them to offer the given amount of employment. It follows, therefore, that, given what we shall call the community's propensity to consume, the equilibrium level of employment, i.e. the level at which there is no inducement to employers as a whole either to expand or to contract employment, will depend on the amount of current investment. The amount of current investment will depend, in turn, on what we shall call the inducement to invest; and the inducement to invest will

be found to depend on the relation between the schedule of the marginal efficiency of capital and the complex of rates of interest on loans of various maturities and risks.

Thus, given the propensity to consume and the rate of new investment, there will be only one level of employment consistent with equilibrium; since any other level will lead to inequality between the aggregate supply price of output as a whole and its aggregate demand price. This level cannot be *greater* than full employment, i.e. the real wage cannot be less than the marginal disutility of labour. But there is no reason in general for expecting it to be *equal* to full employment. The effective demand associated with full employment is a special case, only realised when the propensity to consume and the inducement to invest stand in a particular relationship to one another. This particular relationship, which corresponds to the assumptions of the classical theory, is in a sense an optimum relationship. But it can only exist when, by accident or design, current investment provides an amount of demand just equal to the excess of the aggregate supply price of the output resulting from full employment over what the community will choose to spend on consumption when it is fully employed.

This theory can be summed up in the following propositions:

(1) In a given situation of technique, resources and costs, income (both money-income and real income) depends on the volume of employment N.

(2) The relationship between the community's income and what it can be expected to spend on consumption, designated by D_1, will depend on the psychological characteristic of the community, which we shall call its *propensity to consume*. That is to say, consumption will depend on the level of aggregate income and, therefore, on the level of employment N, except when there is some change in the propensity to consume.

(3) The amount of labour N which the entrepreneurs decide to employ depends on the sum (D) of *two* quantities, namely D_1, the amount which the community is expected to spend on consumption, and D_2, the amount which it is expected to devote to new investment. D is what we have called above the *effective demand*.

(4) Since $D_1 + D_2 = D = \phi(N)$, where ϕ is the aggregate supply function, and since, as we have seen in (2) above, D_1 is a function of N, which we may write $\chi(N)$, depending on the propensity to consume, it follows that $\phi(N) - \chi(N) = D_2$.

(5) Hence the volume of employment in equilibrium depends on (i) the aggregate supply function, ϕ, (ii) the propensity to consume, χ, and (iii) the volume of investment, D_2. This is the essence of the General Theory of Employment.

(6) For every value of N there is a corresponding marginal productivity of labour in the wage-goods industries; and it is this which determines the real wage. (5) is, therefore, subject to the condition that N cannot *exceed* the value which reduces the real wage to equality with the marginal disutility of labour. This means that not all changes in D are compatible with our temporary assumption that money-wages are constant. Thus it will be essential to a full statement of our theory to dispense with this assumption.

(7) On the classical theory, according to which $D = \phi(N)$ for *all* values of N, the volume of employment is in neutral equilibrium for all values of N less than its maximum value; so that the forces of competition between entrepreneurs may be expected to push it to this maximum value. Only at this point, on the classical theory, can there be stable equilibrium.

(8) *When employment increases, D_1 will increase, but not by so much as D;* since when our income increases our consumption increases also, but not by so much. The key to our practical problem is to be found in this

psychological law. For it follows from this that the greater the volume of employment the greater will be the gap between the aggregate supply price (Z) of the corresponding output and the sum (D_1) which the entrepreneurs can expect to get back out of the expenditure of consumers. Hence, if there is no change in the propensity to consume, employment cannot increase, unless at the same time D_2 is increasing so as to fill the increasing gap between Z and D_1. Thus—except on the special assumptions of the classical theory according to which there is some force in operation which, when employment increases, always causes D_2 to increase sufficiently to fill the widening gap between Z and D_1—the economic system may find itself in stable equilibrium with N at a level below full employment, namely at the level given by the intersection of the aggregate demand function with the aggregate supply function.

Thus the volume of employment is not determined by the marginal disutility of labour measured in terms of real wages, except in so far as the supply of labour available at a given real wage sets a *maximum* level to employment. The propensity to consume and the rate of new investment determine between them the volume of employment, and the volume of employment is uniquely related to a given level of real wages—not the other way round. If the propensity to consume and the rate of new investment result in a deficient effective demand, the actual level of employment will fall short of the supply of labour potentially available at the existing real wage, and the equilibrium real wage will be *greater* than the marginal disutility of the equilibrium level of employment.

This analysis supplies us with an explanation of the paradox of poverty in the midst of plenty. For the mere existence of an insufficiency of effective demand may, and often will, bring the increase of employment to a standstill *before* a level of full employ-

ment has been reached. The insufficiency of effective demand will inhibit the process of production in spite of the fact that the marginal product of labour still exceeds in value the marginal disutility of employment.

Moreover the richer the community, the wider will tend to be the gap between its actual and its potential production; and therefore the more obvious and outrageous the defects of the economic system. For a poor community will be prone to consume by far the greater part of its output, so that a very modest measure of investment will be sufficient to provide full employment; whereas a wealthy community will have to discover much ampler opportunities for investment if the saving propensities of its wealthier members are to be compatible with the employment of its poorer members. If in a potentially wealthy community the inducement to invest is weak, then, in spite of its potential wealth, the working of the principle of effective demand will compel it to reduce its actual output, until, in spite of its potential wealth, it has become so poor that its surplus over its consumption is sufficiently diminished to correspond to the weakness of the inducement to invest.

But worse still. Not only is the marginal propensity to consume[1] weaker in a wealthy community, but, owing to its accumulation of capital being already larger, the opportunities for further investment are less attractive unless the rate of interest falls at a sufficiently rapid rate; which brings us to the theory of the rate of interest and to the reasons why it does not automatically fall to the appropriate level, which will occupy Book IV.

Thus the analysis of the propensity to consume, the definition of the marginal efficiency of capital and the theory of the rate of interest are the three main gaps in our existing knowledge which it will be necessary to fill. When this has been accomplished,

[1] Defined in chapter 10, below.

we shall find that the theory of prices falls into its proper place as a matter which is subsidiary to our general theory. We shall discover, however, that money plays an essential part in our theory of the rate of interest; and we shall attempt to disentangle the peculiar characteristics of money which distinguish it from other things.

III

The idea that we can safely neglect the aggregate demand function is fundamental to the Ricardian economics, which underlie what we have been taught for more than a century. Malthus, indeed, had vehemently opposed Ricardo's doctrine that it was impossible for effective demand to be deficient; but vainly. For, since Malthus was unable to explain clearly (apart from an appeal to the facts of common observation) how and why effective demand could be deficient or excessive, he failed to furnish an alternative construction; and Ricardo conquered England as completely as the Holy Inquisition conquered Spain. Not only was his theory accepted by the city, by statesmen and by the academic world. But controversy ceased; the other point of view completely disappeared; it ceased to be discussed. The great puzzle of effective demand with which Malthus had wrestled vanished from economic literature. You will not find it mentioned even once in the whole works of Marshall, Edgeworth and Professor Pigou, from whose hands the classical theory has received its most mature embodiment. It could only live on furtively, below the surface, in the underworlds of Karl Marx, Silvio Gesell or Major Douglas.

The completeness of the Ricardian victory is something of a curiosity and a mystery. It must have been due to a complex of suitabilities in the doctrine to the environment into which it was projected. That it

reached conclusions quite different from what the ordinary uninstructed person would expect, added, I suppose, to its intellectual prestige. That its teaching, translated into practice, was austere and often unpalatable, lent it virtue. That it was adapted to carry a vast and consistent logical superstructure, gave it beauty. That it could explain much social injustice and apparent cruelty as an inevitable incident in the scheme of progress, and the attempt to change such things as likely on the whole to do more harm than good, commended it to authority. That it afforded a measure of justification to the free activities of the individual capitalist, attracted to it the support of the dominant social force behind authority.

But although the doctrine itself has remained unquestioned by orthodox economists up to a late date, its signal failure for purposes of scientific prediction has greatly impaired, in the course of time, the prestige of its practitioners. For professional economists, after Malthus, were apparently unmoved by the lack of correspondence between the results of their theory and the facts of observation;—a discrepancy which the ordinary man has not failed to observe, with the result of his growing unwillingness to accord to economists that measure of respect which he gives to other groups of scientists whose theoretical results are confirmed by observation when they are applied to the facts.

The celebrated *optimism* of traditional economic theory, which has led to economists being looked upon as Candides, who, having left this world for the cultivation of their gardens, teach that all is for the best in the best of all possible worlds provided we will let well alone, is also to be traced, I think, to their having neglected to take account of the drag on prosperity which can be exercised by an insufficiency of effective demand. For there would obviously be a natural tendency towards the optimum employment of resources in a society which was functioning after the

changes in the rate of interest and in fiscal policy may make some difference; but the other objective factors which might affect it, whilst they must not be overlooked, are not likely to be important in ordinary circumstances.

The fact that, given the general economic situation, the expenditure on consumption in terms of the wage-unit depends in the main, on the volume of output and employment is the justification for summing up the other factors in the portmanteau function 'propensity to consume'. For whilst the other factors are capable of varying (and this must not be forgotten), the aggregate income measured in terms of the wage-unit is, as a rule, the principal variable upon which the consumption-constituent of the aggregate demand function will depend.

III

Granted, then, that the propensity to consume is a fairly stable function so that, as a rule, the amount of aggregate consumption mainly depends on the amount of aggregate income (both measured in terms of wage-units), changes in the propensity itself being treated as a secondary influence, what is the normal shape of this function?

The fundamental psychological law, upon which we are entitled to depend with great confidence both *a priori* from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as their income increases, but not by as much as the increase in their income. That is to say, if C_w is the amount of consumption and Y_w is income (both measured in wage-units) ΔC_w has the same sign as ΔY_w but is smaller in amount, i.e. $\frac{dC_w}{dY_w}$ is positive and less than unity.

This is especially the case where we have short periods in view, as in the case of the so-called cyclical fluctuations of employment during which habits, as distinct from more permanent psychological propensities, are not given time enough to adapt themselves to changed objective circumstances. For a man's habitual standard of life usually has the first claim on his income, and he is apt to save the difference which discovers itself between his actual income and the expense of his habitual standard; or, if he does adjust his expenditure to changes in his income, he will over short periods do so imperfectly. Thus a rising income will often be accompanied by increased saving, and a falling income by decreased saving, on a greater scale at first than subsequently.

But, apart from short-period *changes* in the level of income, it is also obvious that a higher absolute level of income will tend, as a rule, to widen the gap between income and consumption. For the satisfaction of the immediate primary needs of a man and his family is usually a stronger motive than the motives towards accumulation, which only acquire effective sway when a margin of comfort has been attained. These reasons will lead, as a rule, to a *greater proportion* of income being saved as real income increases. But whether or not a greater proportion is saved, we take it as a fundamental psychological rule of any modern community that, when its real income is increased, it will not increase its consumption by an equal *absolute* amount, so that a greater absolute amount must be saved, unless a large and unusual change is occurring at the same time in other factors. As we shall show subsequently,[1] the stability of the economic system essentially depends on this rule prevailing in practice. This means that, if employment and hence aggregate income increase, *not all* the additional employment will be required to satisfy the needs of additional consumption.

[1] Cf. p. 251 below.

On the other hand, a decline in income due to a decline in the level of employment, if it goes far, may even cause consumption to exceed income not only by some individuals and institutions using up the financial reserves which they have accumulated in better times, but also by the government, which will be liable, willingly or unwillingly, to run into a budgetary deficit or will provide unemployment relief, for example, out of borrowed money. Thus, when employment falls to a low level, aggregate consumption will decline by a smaller amount than that by which real income has declined, by reason both of the habitual behaviour of individuals and also of the probable policy of governments; which is the explanation why a new position of equilibrium can usually be reached within a modest range of fluctuation. Otherwise a fall in employment and income, once started, might proceed to extreme lengths.

This simple principle leads, it will be seen, to the same conclusion as before, namely, that employment can only increase *pari passu* with an increase in investment; unless, indeed, there is a change in the propensity to consume. For since consumers will spend less than the increase in aggregate supply price when employment is increased, the increased employment will prove unprofitable unless there is an increase in investment to fill the gap.

IV

We must not underestimate the importance of the fact already mentioned above that, whereas employment is a function of the expected consumption and the expected investment, consumption is, *cet. par.*, a function of *net* income, i.e. of *net* investment (net income being equal to consumption *plus* net investment). In other words, the larger the financial provision which it is thought necessary to make before reckoning net in-

come, the less favourable to consumption, and therefore to employment, will a given level of investment prove to be.

When the whole of this financial provision (or supplementary cost) is in fact currently expended in the upkeep of the already existing capital equipment, this point is not likely to be overlooked. But when the financial provision *exceeds* the actual expenditure on current upkeep, the practical results of this in its effect on employment are not always appreciated. For the amount of this excess neither directly gives rise to current investment nor is available to pay for consumption. It has, therefore, to be balanced by new investment, the demand for which has arisen quite independently of the current wastage of old equipment against which the financial provision is being made; with the result that the new investment available to provide current income is correspondingly diminished and a more intense demand for new investment is necessary to make possible a given level of employment. Moreover, much the same considerations apply to the allowance for wastage included in user cost, in so far as the wastage is not actually made good.

Take a house which continues to be habitable until it is demolished or abandoned. If a certain sum is written off its value out of the annual rent paid by the tenants, which the landlord neither spends on upkeep nor regards as net income available for consumption, this provision, whether it is a part of U or of V, constitutes a drag on employment all through the life of the house, suddenly made good in a lump when the house has to be rebuilt.

In a stationary economy all this might not be worth mentioning, since in each year the depreciation allowances in respect of old houses would be exactly offset by the new houses built in replacement of those reaching the end of their lives in that year. But such factors may be serious in a non-static economy, especially

depends on how far the rate of interest is favourable to investment, after taking account of the marginal efficiency of capital.[1] No, this is an overstatement. If the rate of interest were so governed as to maintain continuous full employment, virtue would resume her sway;—the rate of capital accumulation would depend on the weakness of the propensity to consume. Thus, once again, the tribute that classical economists pay to her is due to their concealed assumption that the rate of interest always is so governed.

[1] In some passages of this section we have tacitly anticipated ideas which will be introduced in Book IV.

Chapter 10

THE MARGINAL PROPENSITY TO CONSUME AND THE MULTIPLIER

We established in chapter 8 that employment can only increase *pari passu* with investment unless there is a change in the propensity to consume. We can now carry this line of thought a stage further. For in given circumstances a definite ratio, to be called the *multiplier*, can be established between income and investment and, subject to certain simplifications, between the total employment and the employment directly employed on investment (which we shall call the *primary employment*). This further step is an integral part of our theory of employment, since it establishes a precise relationship, given the propensity to consume, between aggregate employment and income and the rate of investment. The conception of the multiplier was first introduced into economic theory by Mr R. F. Kahn in his article on 'The Relation of Home Investment to Unemployment' (*Economic Journal*, June 1931). His argument in this article depended on the fundamental notion that, if the propensity to consume in various hypothetical circumstances is (together with certain other conditions) taken as given and we conceive the monetary or other public authority to take steps to stimulate or to retard investment, the change in the amount of employment will be a function of the net change in the amount of investment; and it aimed at laying down general principles by which to estimate the actual quantitative relationship between an incre-

ment of net investment and the increment of aggregate employment which will be associated with it. Before coming to the multiplier, however, it will be convenient to introduce the conception of the *marginal propensity to consume*.

I

The fluctuations in real income under consideration in this book are those which result from applying different quantities of employment (i.e. of labour-units) to a given capital equipment, so that real income increases and decreases with the number of labour-units employed. If, as we assume in general, there is a decreasing return at the margin as the number of labour-units employed on the given capital equipment is increased, income measured in terms of wage-units will increase more than in proportion to the amount of employment, which, in turn, will increase more than in proportion to the amount of real income measured (if that is possible) in terms of product. Real income measured in terms of product and income measured in terms of wage-units will, however, increase and decrease together (in the short period when capital equipment is virtually unchanged). Since, therefore, real income, in terms of product, may be incapable of precise numerical measurement, it is often convenient to regard income in terms of wage-units (Y_w) as an adequate working index of changes in real income. In certain contexts we must not overlook the fact that, in general, Y_w increases and decreases in a greater proportion than real income; but in other contexts the fact that they always increase and decrease together renders them virtually interchangeable.

Our normal psychological law that, when the real income of the community increases or decreases, its consumption will increase or decrease but not so fast, can, therefore, be translated—not, indeed, with absolute accuracy but subject to qualifications which are obvious

and can easily be stated in a formally complete fashion—into the propositions that ΔC_w and ΔY_w have the same sign, but $\Delta Y_w > \Delta C_w$, where C_w is the consumption in terms of wage-units. This is merely a repetition of the proposition already established on p. 29 above. Let us define, then, $\frac{dC_w}{dY_w}$ as the *marginal propensity to consume*.

This quantity is of considerable importance, because it tells us how the next increment of output will have to be divided between consumption and investment. For $\Delta Y_w = \Delta C_w + \Delta I_w$, where ΔC_w and ΔI_w are the increments of consumption and investment; so that we can write $\Delta Y_w = k\Delta I_w$, where $1 - \frac{1}{k}$ is equal to the marginal propensity to consume.

Let us call k the *investment multiplier*. It tells us that, when there is an increment of aggregate investment, income will increase by an amount which is k times the increment of investment.

II

Mr Kahn's multiplier is a little different from this, being what we may call the *employment multiplier* designated by k', since it measures the ratio of the increment of total employment which is associated with a given increment of primary employment in the investment industries. That is to say, if the increment of investment ΔI_w leads to an increment of primary employment ΔN_2 in the investment industries, the increment of total employment $\Delta N = k'\Delta N_2$.

There is no reason in general to suppose that $k = k'$. For there is no necessary presumption that the shapes of the relevant portions of the aggregate supply functions for different types of industry are such that the ratio of the increment of employment in the one set of industries to the increment of demand which has

stimulated it will be the same as in the other set of industries.[1] It is easy, indeed, to conceive of cases, as, for example, where the marginal propensity to consume is widely different from the average propensity, in which there would be a presumption in favour of some inequality between $\frac{\Delta Y_w}{\Delta N}$ and $\frac{\Delta I_w}{\Delta N_2}$, since there would be very divergent proportionate changes in the demands for consumption-goods and investment-goods respectively. If we wish to take account of such possible differences in the shapes of the relevant portions of the aggregate supply functions for the two groups of industries respectively, there is no difficulty in rewriting the following argument in the more generalised form. But to elucidate the ideas involved, it will be convenient to deal with the simplified case where $k = k'$.

It follows, therefore, that, if the consumption psychology of the community is such that they will choose to consume, e.g. nine-tenths of an increment of income,[2] then the multiplier k is 10; and the total employment caused by (e.g.) increased public works will be ten times the primary employment provided by

[1] More precisely, if e_e and e'_e are the elasticities of employment in industry as a whole and in the investment industries respectively, and if N and N_2 are the numbers of men employed in industry as a whole and in the investment industries, we have

$$\Delta Y_w = \frac{Y_w}{e_e . N}\Delta N$$

and

$$\Delta I_w = \frac{I_w}{e'_e . N_2}\Delta N_2,$$

so that

$$\Delta N = \frac{e_e I_w N}{e'_e N_2 Y_w} k . \Delta N_2,$$

i.e.

$$k' = \frac{I_w}{e'_e N_2} . \frac{e_e N}{Y_w} k.$$

If, however, there is no reason to expect any material relevant difference in the shapes of the aggregate supply functions for industry as a whole and for the investment industries respectively, so that $\frac{I_w}{e'_e . N_2} = \frac{Y_w}{e_e . N}$, then it follows that $\frac{\Delta Y_w}{\Delta N} = \frac{\Delta I_w}{\Delta N_2}$ and, therefore, that $k = k'$.

[2] Our quantities are measured throughout in terms of wage-units.

the public works themselves, assuming no reduction of investment in other directions. Only in the event of the community maintaining their consumption unchanged in spite of the increase in employment and hence in real income, will the increase of employment be restricted to the primary employment provided by the public works. If, on the other hand, they seek to consume the whole of any increment of income, there will be no point of stability and prices will rise without limit. With normal psychological suppositions, an increase in employment will only be associated with a decline in consumption if there is at the same time a change in the propensity to consume—as the result, for instance, of propaganda in time of war in favour of restricting individual consumption; and it is only in this event that the increased employment in investment will be associated with an unfavourable repercussion on employment in the industries producing for consumption.

This only sums up in a formula what should by now be obvious to the reader on general grounds. An increment of investment in terms of wage-units cannot occur unless the public are prepared to increase their savings in terms of wage-units. Ordinarily speaking, the public will not do this unless their aggregate income in terms of wage-units is increasing. Thus their effort to consume a part of their increased incomes will stimulate output until the new level (and distribution) of incomes provides a margin of saving sufficient to correspond to the increased investment. The multiplier tells us by how much their employment has to be increased to yield an increase in real income sufficient to induce them to do the necessary extra saving, and is a function of their psychological propensities.[1] If saving is the pill and consumption is the jam, the extra jam has to be proportioned to the size of the

[1] Though in the more generalised case it is also a function of the physical conditions of production in the investment and consumption industries respectively.

additional pill. Unless the psychological propensities of the public are different from what we are supposing, we have here established the law that increased employment for investment must necessarily stimulate the industries producing for consumption and thus lead to a total increase of employment which is a multiple of the primary employment required by the investment itself.

It follows from the above that, if the marginal propensity to consume is not far short of unity, small fluctuations in investment will lead to wide fluctuations in employment; but, at the same time, a comparatively small increment of investment will lead to full employment. If, on the other hand, the marginal propensity to consume is not much above zero, small fluctuations in investment will lead to correspondingly small fluctuations in employment; but, at the same time, it may require a large increment of investment to produce full employment. In the former case involuntary unemployment would be an easily remedied malady, though liable to be troublesome if it is allowed to develop. In the latter case, employment may be less variable but liable to settle down at a low level and to prove recalcitrant to any but the most drastic remedies. In actual fact the marginal propensity to consume seems to lie somewhere between these two extremes, though much nearer to unity than to zero; with the result that we have, in a sense, the worst of both worlds, fluctuations in employment being considerable and, at the same time, the increment in investment required to produce full employment being too great to be easily handled. Unfortunately the fluctuations have been sufficient to prevent the nature of the malady from being obvious, whilst its severity is such that it cannot be remedied unless its nature is understood.

When full employment is reached, any attempt to increase investment still further will set up a tendency in money-prices to rise without limit, irrespective of the marginal propensity to consume; i.e. we shall

have reached a state of true inflation.[1] Up to this point, however, rising prices will be associated with an increasing aggregate real income.

III

We have been dealing so far with a *net* increment of investment. If, therefore, we wish to apply the above without qualification to the effect of (e.g.) increased public works, we have to assume that there is no offset through decreased investment in other directions,—and also, of course, no associated change in the propensity of the community to consume. Mr Kahn was mainly concerned in the article referred to above in considering what offsets we ought to take into account as likely to be important, and in suggesting quantitative estimates. For in an actual case there are several factors besides some specific increase of investment of a given kind which enter into the final result. If, for example, a government employs 100,000 additional men on public works, and if the multiplier (as defined above) is 4, it is not safe to assume that aggregate employment will increase by 400,000. For the new policy may have adverse reactions on investment in other directions.

It would seem (following Mr Kahn) that the following are likely in a modern community to be the factors which it is most important not to overlook (though the first two will not be fully intelligible until after Book IV has been reached):

(i) The method of financing the policy and the increased working cash, required by the increased employment and the associated rise of prices, may have the effect of increasing the rate of interest and so retarding investment in other directions, unless the monetary authority takes steps to the contrary; whilst, at the same time, the increased cost of capital goods will reduce their marginal efficiency to the private in-

[1] Cf. chapter 21, p. 303, below.

Chapter 11

THE MARGINAL EFFICIENCY OF CAPITAL

I

When a man buys an investment or capital-asset, he purchases the right to the series of prospective returns, which he expects to obtain from selling its output, after deducting the running expenses of obtaining that output, during the life of the asset. This series of annuities Q_1, Q_2, ... Q_n it is convenient to call the *prospective yield* of the investment.

Over against the prospective yield of the investment we have the *supply price* of the capital-asset, meaning by this, not the market-price at which an asset of the type in question can actually be purchased in the market, but the price which would just induce a manufacturer newly to produce an additional unit of such assets, i.e. what is sometimes called its *replacement cost*. The relation between the prospective yield of a capital-asset and its supply price or replacement cost, i.e. the relation between the prospective yield of one more unit of that type of capital and the cost of producing that unit, furnishes us with the *marginal efficiency of capital* of that type. More precisely, I define the marginal efficiency of capital as being equal to that rate of discount which would make the present value of the series of annuities given by the returns expected from the capital-asset during its life just equal to its supply price. This gives us the marginal efficiencies of particular types of capital-assets. The greatest of

these marginal efficiencies can then be regarded as the marginal efficiency of capital in general.

The reader should note that the marginal efficiency of capital is here defined in terms of the *expectation* of yield and of the *current* supply price of the capital-asset. It depends on the rate of return expected to be obtainable on money if it were invested in a *newly* produced asset; not on the historical result of what an investment has yielded on its original cost if we look back on its record after its life is over.

If there is an increased investment in any given type of capital during any period of time, the marginal efficiency of that type of capital will diminish as the investment in it is increased, partly because the prospective yield will fall as the supply of that type of capital is increased, and partly because, as a rule, pressure on the facilities for producing that type of capital will cause its supply price to increase; the second of these factors being usually the more important in producing equilibrium in the short run, but the longer the period in view the more does the first factor take its place. Thus for each type of capital we can build up a schedule, showing by how much investment in it will have to increase within the period, in order that its marginal efficiency should fall to any given figure. We can then aggregate these schedules for all the different types of capital, so as to provide a schedule relating the rate of aggregate investment to the corresponding marginal efficiency of capital in general which that rate of investment will establish. We shall call this the investment demand-schedule; or, alternatively, the schedule of the marginal efficiency of capital.

Now it is obvious that the actual rate of current investment will be pushed to the point where there is no longer any class of capital-asset of which the marginal efficiency exceeds the current rate of interest. In other words, the rate of investment will be pushed to the

point on the investment demand-schedule where the marginal efficiency of capital in general is equal to the market rate of interest.[1]

The same thing can also be expressed as follows. If Q_r is the prospective yield from an asset at time r, and d_r is the present value of £1 deferred r years *at the current rate of interest*, $\Sigma Q_r d_r$ is the demand price of the investment; and investment will be carried to the point where $\Sigma Q_r d_r$ becomes equal to the supply price of the investment as defined above. If, on the other hand, $\Sigma Q_r d_r$ falls short of the supply price, there will be no current investment in the asset in question.

It follows that the inducement to invest depends partly on the investment demand-schedule and partly on the rate of interest. Only at the conclusion of Book IV will it be possible to take a comprehensive view of the factors determining the rate of investment in their actual complexity. I would, however, ask the reader to note at once that neither the knowledge of an asset's prospective yield nor the knowledge of the marginal efficiency of the asset enables us to deduce either the rate of interest or the present value of the asset. We must ascertain the rate of interest from some other source, and only then can we value the asset by 'capitalising' its prospective yield

II

How is the above definition of the marginal efficiency of capital related to common usage? The *Marginal Productivity* or *Yield* or *Efficiency* or *Utility* of Capital are familiar terms which we have all frequently used. But it is not easy by searching the literature of economics to

[1] For the sake of simplicity of statement I have slurred the point that we are dealing with complexes of rates of interest and discount corresponding to the different lengths of time which will elapse before the various prospective returns from the asset are realised. But it is not difficult to re-state the argument so as to cover this point.

current phenomenon; and if we reduce the marginal efficiency of capital to the same status, we cut ourselves off from taking any direct account of the influence of the future in our analysis of the existing equilibrium.

The fact that the assumptions of the static state often underlie present-day economic theory, imports into it a large element of unreality. But the introduction of the concepts of user cost and of the marginal efficiency of capital, as defined above, will have the effect, I think, of bringing it back to reality, whilst reducing to a minimum the necessary degree of adaptation.

It is by reason of the existence of durable equipment that the economic future is linked to the present. It is, therefore, consonant with, and agreeable to, our broad principles of thought, that the expectation of the future should affect the present through the demand price for durable equipment.

Chapter 12

THE STATE OF LONG-TERM EXPECTATION

I

We have seen in the previous chapter that the scale of investment depends on the relation between the rate of interest and the schedule of the marginal efficiency of capital corresponding to different scales of current investment, whilst the marginal efficiency of capital depends on the relation between the supply price of a capital-asset and its prospective yield. In this chapter we shall consider in more detail some of the factors which determine the prospective yield of an asset.

The considerations upon which expectations of prospective yields are based are partly existing facts which we can assume to be known more or less for certain, and partly future events which can only be forecasted with more or less confidence. Amongst the first may be mentioned the existing stock of various types of capital-assets and of capital-assets in general and the strength of the existing consumers' demand for goods which require for their efficient production a relatively larger assistance from capital. Amongst the latter are future changes in the type and quantity of the stock of capital-assets and in the tastes of the consumer, the strength of effective demand from time to time during the life of the investment under consideration, and the changes in the wage-unit in terms of money which may occur during its life. We may sum up the state of psychological expectation which covers the

latter as being the *state of long-term expectation*;—as distinguished from the short-term expectation upon the basis of which a producer estimates what he will get for a product when it is finished if he decides to begin producing it to-day with the existing plant, which we examined in chapter 5.

II

It would be foolish, in forming our expectations, to attach great weight to matters which are very uncertain.[1] It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and scanty. For this reason the facts of the existing situation enter, in a sense disproportionately, into the formation of our long-term expectations; our usual practice being to take the existing situation and to project it into the future, modified only to the extent that we have more or less definite reasons for expecting a change.

The state of long-term expectation, upon which our decisions are based, does not solely depend, therefore, on the most probable forecast we can make. It also depends on the *confidence* with which we make this forecast—on how highly we rate the likelihood of our best forecast turning out quite wrong. If we expect large changes but are very uncertain as to what precise form these changes will take, then our confidence will be weak.

The *state of confidence*, as they term it, is a matter to which practical men always pay the closest and most anxious attention. But economists have not analysed it carefully and have been content, as a rule, to discuss

[1] By 'very uncertain' I do not mean the same thing as 'very improbable'. Cf. my *Treatise on Probability* [*JMK*, vol. VIII], chap. 6, on 'The Weight of Arguments'.

it in general terms. In particular it has not been made clear that its relevance to economic problems comes in through its important influence on the schedule of the marginal efficiency of capital. There are not two separate factors affecting the rate of investment, namely, the schedule of the marginal efficiency of capital and the state of confidence. The state of confidence is relevant because it is one of the major factors determining the former, which is the same thing as the investment demand-schedule.

There is, however, not much to be said about the state of confidence *a priori*. Our conclusions must mainly depend upon the actual observation of markets and business psychology. This is the reason why the ensuing digression is on a different level of abstraction from most of this book.

For convenience of exposition we shall assume in the following discussion of the state of confidence that there are no changes in the rate of interest; and we shall write, throughout the following sections, as if changes in the values of investments were solely due to changes in the expectation of their prospective yields and not at all to changes in the rate of interest at which these prospective yields are capitalised. The effect of changes in the rate of interest is, however, easily superimposed on the effect of changes in the state of confidence.

III

The outstanding fact is the extreme precariousness of the basis of knowledge on which our estimates of prospective yield have to be made. Our knowledge of the factors which will govern the yield of an investment some years hence is usually very slight and often negligible. If we speak frankly, we have to admit that our basis of knowledge for estimating the yield ten years hence of a railway, a copper mine, a textile factory, the goodwill of a patent medicine, an Atlantic

liner, a building in the City of London amounts to little and sometimes to nothing; or even five years hence. In fact, those who seriously attempt to make any such estimate are often so much in the minority that their behaviour does not govern the market.

In former times, when enterprises were mainly owned by those who undertook them or by their friends and associates, investment depended on a sufficient supply of individuals of sanguine temperament and constructive impulses who embarked on business as a way of life, not really relying on a precise calculation of prospective profit. The affair was partly a lottery, though with the ultimate result largely governed by whether the abilities and character of the managers were above or below the average. Some would fail and some would succeed. But even after the event no one would know whether the average results in terms of the sums invested had exceeded, equalled or fallen short of the prevailing rate of interest; though, if we exclude the exploitation of natural resources and monopolies, it is probable that the actual average results of investments, even during periods of progress and prosperity, have disappointed the hopes which prompted them. Business men play a mixed game of skill and chance, the average results of which to the players are not known by those who take a hand. If human nature felt no temptation to take a chance, no satisfaction (profit apart) in constructing a factory, a railway, a mine or a farm, there might not be much investment merely as a result of cold calculation.

Decisions to invest in private business of the old-fashioned type were, however, decisions largely irrevocable, not only for the community as a whole, but also for the individual. With the separation between ownership and management which prevails to-day and with the development of organised investment markets, a new factor of great importance has entered in, which sometimes facilitates investment but sometimes adds

greatly to the instability of the system. In the absence of security markets, there is no object in frequently attempting to revalue an investment to which we are committed. But the Stock Exchange revalues many investments every day and the revaluations give a frequent opportunity to the individual (though not to the community as a whole) to revise his commitments. It is as though a farmer, having tapped his barometer after breakfast, could decide to remove his capital from the farming business between 10 and 11 in the morning and reconsider whether he should return to it later in the week. But the daily revaluations of the Stock Exchange, though they are primarily made to facilitate transfers of old investments between one individual and another, inevitably exert a decisive influence on the rate of current investment. For there is no sense in building up a new enterprise at a cost greater than that at which a similar existing enterprise can be purchased; whilst there is an inducement to spend on a new project what may seem an extravagant sum, if it can be floated off on the Stock Exchange at an immediate profit.[1] Thus certain classes of investment are governed by the average expectation of those who deal on the Stock Exchange as revealed in the price of shares, rather than by the genuine expectations of the professional entrepreneur.[2] How then are these highly significant daily, even hourly, revaluations of existing investments carried out in practice?

[1] In my *Treatise on Money* (vol. ii. p. 195) [*JMK*, vol. vi, p. 174] I pointed out that when a company's shares are quoted very high so that it can raise more capital by issuing more shares on favourable terms, this has the same effect as if it could borrow at a low rate of interest. I should now describe this by saying that a high quotation for existing equities involves an increase in the marginal efficiency of the corresponding type of capital and therefore has the same effect (since investment depends on a comparison between the marginal efficiency of capital and the rate of interest) as a fall in the rate of interest.

[2] This does not apply, of course, to classes of enterprise which are not readily marketable or to which no negotiable instrument closely corresponds. The categories falling within this exception were formerly extensive. But measured as a proportion of the total value of new investment they are rapidly declining in importance.

IV

In practice we have tacitly agreed, as a rule, to fall back on what is, in truth, a *convention*. The essence of this convention—though it does not, of course, work out quite so simply—lies in assuming that the existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change. This does not mean that we really believe that the existing state of affairs will continue indefinitely. We know from extensive experience that this is most unlikely. The actual results of an investment over a long term of years very seldom agree with the initial expectation. Nor can we rationalise our behaviour by arguing that to a man in a state of ignorance errors in either direction are equally probable, so that there remains a mean actuarial expectation based on equi-probabilities. For it can easily be shown that the assumption of arithmetically equal probabilities based on a state of ignorance leads to absurdities. We are assuming, in effect, that the existing market valuation, however arrived at, is uniquely *correct* in relation to our existing knowledge of the facts which will influence the yield of the investment, and that it will only change in proportion to changes in this knowledge; though, philosophically speaking, it cannot be uniquely correct, since our existing knowledge does not provide a sufficient basis for a calculated mathematical expectation. In point of fact, all sorts of considerations enter into the market valuation which are in no way relevant to the prospective yield.

Nevertheless the above conventional method of calculation will be compatible with a considerable measure of continuity and stability in our affairs, *so long as we can rely on the maintenance of the convention.*

For if there exist organised investment markets and if we can rely on the maintenance of the convention, an investor can legitimately encourage himself with the

idea that the only risk he runs is that of a genuine change in the news *over the near future*, as to the likelihood of which he can attempt to form his own judgment, and which is unlikely to be very large. For, assuming that the convention holds good, it is only these changes which can affect the value of his investment, and he need not lose his sleep merely because he has not any notion what his investment will be worth ten years hence. Thus investment becomes reasonably 'safe' for the individual investor over short periods, and hence over a succession of short periods however many, if he can fairly rely on there being no breakdown in the convention and on his therefore having an opportunity to revise his judgment and change his investment, before there has been time for much to happen. Investments which are 'fixed' for the community are thus made 'liquid' for the individual.

It has been, I am sure, on the basis of some such procedure as this that our leading investment markets have been developed. But it is not surprising that a convention, in an absolute view of things so arbitrary, should have its weak points. It is its precariousness which creates no small part of our contemporary problem of securing sufficient investment.

V

Some of the factors which accentuate this precariousness may be briefly mentioned.

(1) As a result of the gradual increase in the proportion of the equity in the community's aggregate capital investment which is owned by persons who do not manage and have no special knowledge of the circumstances, either actual or prospective, of the business in question, the element of real knowledge in the valuation of investments by whose who own them or contemplate purchasing them has seriously declined.

(2) Day-to-day fluctuations in the profits of existing

investments, which are obviously of an ephemeral and non-significant character, tend to have an altogether excessive, and even an absurd, influence on the market. It is said, for example, that the shares of American companies which manufacture ice tend to sell at a higher price in summer when their profits are seasonally high than in winter when no one wants ice. The recurrence of a bank-holiday may raise the market valuation of the British railway system by several million pounds.

(3) A conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield; since there will be no strong roots of conviction to hold it steady. In abnormal times in particular, when the hypothesis of an indefinite continuance of the existing state of affairs is less plausible than usual even though there are no express grounds to anticipate a definite change, the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for a reasonable calculation.

(4) But there is one feature in particular which deserves our attention. It might have been supposed that competition between expert professionals, possessing judgment and knowledge beyond that of the average private investor, would correct the vagaries of the ignorant individual left to himself. It happens, however, that the energies and skill of the professional investor and speculator are mainly occupied otherwise. For most of these persons are, in fact, largely concerned, not with making superior long-term forecasts of the probable yield of an investment over its whole life, but with foreseeing changes in the conventional basis of valuation a short time ahead of the general public. They are concerned, not with what an investment is

really worth to a man who buys it 'for keeps', but with what the market will value it at, under the influence of mass psychology, three months or a year hence. Moreover, this behaviour is not the outcome of a wrong-headed propensity. It is an inevitable result of an investment market organised along the lines described. For it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30, if you also believe that the market will value it at 20 three months hence.

Thus the professional investor is forced to concern himself with the anticipation of impending changes, in the news or in the atmosphere, of the kind by which experience shows that the mass psychology of the market is most influenced. This is the inevitable result of investment markets organised with a view to so-called 'liquidity'. Of the maxims of orthodox finance none, surely, is more anti-social than the fetish of liquidity, the doctrine that it is a positive virtue on the part of investment institutions to concentrate their resources upon the holding of 'liquid' securities. It forgets that there is no such thing as liquidity of investment for the community as a whole. The social object of skilled investment should be to defeat the dark forces of time and ignorance which envelop our future. The actual, private object of the most skilled investment to-day is 'to beat the gun', as the Americans so well express it, to outwit the crowd, and to pass the bad, or depreciating, half-crown to the other fellow.

This battle of wits to anticipate the basis of conventional valuation a few months hence, rather than the prospective yield of an investment over a long term of years, does not even require gulls amongst the public to feed the maws of the professional;—it can be played by professionals amongst themselves. Nor is it necessary that anyone should keep his simple faith in the conventional basis of valuation having any genuine long-term validity. For it is, so to speak, a game of Snap,

of Old Maid, of Musical Chairs—a pastime in which he is victor who says *Snap* neither too soon nor too late, who passed the Old Maid to his neighbour before the game is over, who secures a chair for himself when the music stops. These games can be played with zest and enjoyment, though all the players know that it is the Old Maid which is circulating, or that when the music stops some of the players will find themselves unseated.

Or, to change the metaphor slightly, professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.

If the reader interjects that there must surely be large profits to be gained from the other players in the long run by a skilled individual who, unperturbed by the prevailing pastime, continues to purchase investments on the best genuine long-term expectations he can frame, he must be answered, first of all, that there are, indeed, such serious-minded individuals and that it makes a vast difference to an investment market whether or not they predominate in their influence over the game-players. But we must also add that there are

several factors which jeopardise the predominance of such individuals in modern investment markets. Investment based on genuine long-term expectation is so difficult to-day as to be scarcely practicable. He who attempts it must surely lead much more laborious days and run greater risks than he who tries to guess better than the crowd how the crowd will behave; and, given equal intelligence, he may make more disastrous mistakes. There is no clear evidence from experience that the investment policy which is socially advantageous coincides with that which is most profitable. It needs *more* intelligence to defeat the forces of time and our ignorance of the future than to beat the gun. Moreover, life is not long enough;—human nature desires quick results, there is a peculiar zest in making money quickly, and remoter gains are discounted by the average man at a very high rate. The game of professional investment is intolerably boring and overexacting to anyone who is entirely exempt from the gambling instinct; whilst he who has it must pay to this propensity the appropriate toll. Furthermore, an investor who proposes to ignore near-term market fluctuations needs greater resources for safety and must not operate on so large a scale, if at all, with borrowed money—a further reason for the higher return from the pastime to a given stock of intelligence and resources. Finally it is the long-term investor, he who most promotes the public interest, who will in practice come in for most criticism, wherever investment funds are managed by committees or boards or banks.[1] For it is in the essence of his behaviour that he should be eccentric, unconventional and rash in the eyes of average opinion. If he is successful, that will only confirm the general belief in his rashness; and if

[1] The practice, usually considered prudent, by which an investment trust or an insurance office frequently calculates not only the income from its investment portfolio but also its capital valuation in the market, may also tend to direct too much attention to short-term fluctuations in the latter.

in the short run he is unsuccessful, which is very likely, he will not receive much mercy. Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally.

(5) So far we have had chiefly in mind the state of confidence of the speculator or speculative investor himself and may have seemed to be tacitly assuming that, if he himself is satisfied with the prospects, he has unlimited command over money at the market rate of interest. This is, of course, not the case. Thus we must also take account of the other facet of the state of confidence, namely, the confidence of the lending institutions towards those who seek to borrow from them, sometimes described as the state of credit. A collapse in the price of equities, which has had disastrous reactions on the marginal efficiency of capital, may have been due to the weakening either of speculative confidence or of the state of credit. But whereas the weakening of either is enough to cause a collapse, recovery requires the revival of *both*. For whilst the weakening of credit is sufficient to bring about a collapse, its strengthening, though a necessary condition of recovery, is not a sufficient condition.

VI

These considerations should not lie beyond the purview of the economist. But they must be relegated to their right perspective. If I may be allowed to appropriate the term *speculation* for the activity of forecasting the psychology of the market, and the term *enterprise* for the activity of forecasting the prospective yield of assets over their whole life, it is by no means always the case that speculation predominates over enterprise. As the organisation of investment markets improves, the risk of the predominance of speculation does, however, increase. In one of the greatest investment markets in the world, namely, New York, the

influence of speculation (in the above sense) is enormous. Even outside the field of finance, Americans are apt to be unduly interested in discovering what average opinion believes average opinion to be; and this national weakness finds its nemesis in the stock market. It is rare, one is told, for an American to invest, as many Englishmen still do, 'for income'; and he will not readily purchase an investment except in the hope of capital appreciation. This is only another way of saying that, when he purchases an investment, the American is attaching his hopes, not so much to its prospective yield, as to a favourable change in the conventional basis of valuation, i.e. that he is, in the above sense, a speculator. Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation. When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done. The measure of success attained by Wall Street, regarded as an institution of which the proper social purpose is to direct new investment into the most profitable channels in terms of future yield, cannot be claimed as one of the outstanding triumphs of *laissez-faire* capitalism—which is not surprising, if I am right in thinking that the best brains of Wall Street have been in fact directed towards a different object.

These tendencies are a scarcely avoidable outcome of our having successfully organised 'liquid' investment markets. It is usually agreed that casinos should, in the public interest, be inaccessible and expensive. And perhaps the same is true of stock exchanges. That the sins of the London Stock Exchange are less than those of Wall Street may be due, not so much to differences in national character, as to the fact that to the average Englishman Throgmorton Street is, compared with Wall Street to the average American, inaccessible and very expensive. The jobber's 'turn', the high

brokerage charges and the heavy transfer tax payable to the Exchequer, which attend dealings on the London Stock Exchange, sufficiently diminish the liquidity of the market (although the practice of fortnightly accounts operates the other way) to rule out a large proportion of the transaction characteristic of Wall Street.[1] The introduction of a substantial government transfer tax on all transactions might prove the most serviceable reform available, with a view to mitigating the predominance of speculation over enterprise in the United States.

The spectacle of modern investment markets has sometimes moved me towards the conclusion that to make the purchase of an investment permanent and indissoluble, like marriage, except by reason of death or other grave cause, might be a useful remedy for our contemporary evils. For this would force the investor to direct his mind to the long-term prospects and to those only. But a little consideration of this expedient brings us up against a dilemma, and shows us how the liquidity of investment markets often facilitates, though it sometimes impedes, the course of new investment. For the fact that each individual investor flatters himself that his commitment is 'liquid' (though this cannot be true for all investors collectively) calms his nerves and makes him much more willing to run a risk. If individual purchases of investments were rendered illiquid, this might seriously impede new investment, so long as *alternative ways* in which to hold his savings are available to the individual. This is the dilemma. So long as it is open to the individual to employ his wealth in hoarding or lending *money*, the alternative of purchasing actual capital assets cannot be rendered sufficiently attractive (especially to the man who does

[1] It is said that, when Wall Street is active, at least a half of the purchases or sales of investments are entered upon with an intention on the part of the speculator to reverse them *the same day*. This is often true of the commodity exchanges also.

not manage the capital assets and knows very little about them), except by organising markets wherein these assets can be easily realised for money.

The only radical cure for the crises of confidence which afflict the economic life of the modern world would be to allow the individual no choice between consuming his income and ordering the production of the specific capital-asset which, even though it be on precarious evidence, impresses him as the most promising investment available to him. It might be that, at times when he was more than usually assailed by doubts concerning the future, he would turn in his perplexity towards more consumption and less new investment. But that would avoid the disastrous, cumulative and far-reaching repercussions of its being open to him, when thus assailed by doubts, to spend his income neither on the one nor on the other.

Those who have emphasised the social dangers of the hoarding of money have, of course, had something similar to the above in mind. But they have overlooked the possibility that the phenomenon can occur without any change, or at least any commensurate change, in the hoarding of money

VII

Even apart from the instability due to speculation, there is the instability due to the characteristic of human nature that a large proportion of our positive activities depend on spontaneous optimism rather than on a mathematical expectation, whether moral or hedonistic or economic. Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits—of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities. Enterprise

only pretends to itself to be mainly actuated by the statements in its own prospectus, however candid and sincere. Only a little more than an expedition to the South Pole, is it based on an exact calculation of benefits to come. Thus if the animal spirits are dimmed and the spontaneous optimism falters, leaving us to depend on nothing but a mathematical expectation, enterprise will fade and die;—though fears of loss may have a basis no more reasonable than hopes of profit had before.

It is safe to say that enterprise which depends on hopes stretching into the future benefits the community as a whole. But individual initiative will only be adequate when reasonable calculation is supplemented and supported by animal spirits, so that the thought of ultimate loss which often overtakes pioneers, as experience undoubtedly tells us and them, is put aside as a healthy man puts aside the expectation of death.

This means, unfortunately, not only that slumps and depressions are exaggerated in degree, but that economic prosperity is excessively dependent on a political and social atmosphere which is congenial to the average business man. If the fear of a Labour Government or a New Deal depresses enterprise, this need not be the result either of a reasonable calculation or of a plot with political intent;—it is the mere consequence of upsetting the delicate balance of spontaneous optimism. In estimating the prospects of investment, we must have regard, therefore, to the nerves and hysteria and even the digestions and reactions to the weather of those upon whose spontaneous activity it largely depends.

We should not conclude from this that everything depends on waves of irrational psychology. On the contrary, the state of long-term expectation is often steady, and, even when it is not, the other factors exert their compensating effects. We are merely reminding ourselves that human decisions affecting the future, whether personal or political or economic, cannot

depend on strict mathematical expectation, since the basis for making such calculations does not exist; and that it is our innate urge to activity which makes the wheels go round, our rational selves choosing between the alternatives as best we are able, calculating where we can, but often falling back for our motive on whim or sentiment or chance.

VIII

There are, moreover, certain important factors which somewhat mitigate in practice the effects of our ignorance of the future. Owing to the operation of compound interest combined with the likelihood of obsolescence with the passage of time, there are many individual investments of which the prospective yield is legitimately dominated by the returns of the comparatively near future. In the case of the most important class of very long-term investments, namely buildings, the risk can be frequently transferred from the investor to the occupier, or at least shared between them, by means of long-term contracts, the risk being outweighed in the mind of the occupier by the advantages of continuity and security of tenure. In the case of another important class of long-term investments, namely public utilities, a substantial proportion of the prospective yield is practically guaranteed by monopoly privileges coupled with the right to charge such rates as will provide a certain stipulated margin. Finally there is a growing class of investments entered upon by, or at the risk of, public authorities, which are frankly influenced in making the investment by a general presumption of there being prospective social advantages from the investment, whatever its commercial yield may prove to be within a wide range, and without seeking to be satisfied that the mathematical expectation of the yield is at least equal to the current rate of interest,—though the rate which the public

authority has to pay may still play a decisive part in determining the scale of investment operations which it can afford.

Thus after giving full weight to the importance of the influence of short-period changes in the state of long-term expectation as distinct from changes in the rate of interest, we are still entitled to return to the latter as exercising, at any rate, in normal circumstances, a great, though not a decisive, influence on the rate of investment. Only experience, however, can show how far management of the rate of interest is capable of continuously stimulating the appropriate volume of investment.

For my own part I am now somewhat sceptical of the success of a merely monetary policy directed towards influencing the rate of interest. I expect to see the State, which is in a position to calculate the marginal efficiency of capital-goods on long views and on the basis of the general social advantage, taking an ever greater responsibility for directly organising investment; since it seems likely that the fluctuations in the market estimation of the marginal efficiency of different types of capital, calculated on the principles I have described above, will be too great to be offset by any practicable changes in the rate of interest.

Chapter 13

THE GENERAL THEORY OF THE RATE OF INTEREST

I

We have shown in chapter 11 that, whilst there are forces causing the rate of investment to rise or fall so as to keep the marginal efficiency of capital equal to the rate of interest, yet the marginal efficiency of capital is, in itself, a different thing from the ruling rate of interest. The schedule of the marginal efficiency of capital may be said to govern the terms on which loanable funds are demanded for the purpose of new investment; whilst the rate of interest governs the terms on which funds are being currently supplied. To complete our theory, therefore, we need to know what determines the rate of interest.

In chapter 14 and its Appendix we shall consider the answers to this question which have been given hitherto. Broadly speaking, we shall find that they make the rate of interest to depend on the interaction of the schedule of the marginal efficiency of capital with the psychological propensity to save. But the notion that the rate of interest is the balancing factor which brings the demand for saving in the shape of new investment forthcoming at a given rate of interest into equality with the supply of saving which results at that rate of interest from the community's psychological propensity to save, breaks down as soon as we perceive that it is impossible to deduce the rate of interest merely from a knowledge of these two factors.

What, then, is our own answer to this question?

II

The psychological time-preferences of an individual require two distinct sets of decisions to carry them out completely. The first is concerned with that aspect of time-preference which I have called the *propensity to consume*, which, operating under the influence of the various motives set forth in Book III, determines for each individual how much of his income he will consume and how much he will reserve in *some* form of command over future consumption.

But this decision having been made, there is a further decision which awaits him, namely, in *what form* he will hold the command over future consumption which he has reserved, whether out of his current income or from previous savings. Does he want to hold it in the form of immediate, liquid command (i.e. in money or its equivalent)? Or is he prepared to part with immediate command for a specified or indefinite period, leaving it to future market conditions to determine on what terms he can, if necessary, convert deferred command over specific goods into immediate command over goods in general? In other words, what is the degree of his *liquidity-preference*—where an individual's liquidity-preference is given by a schedule of the amounts of his resources, valued in terms of money or of wage-units, which he will wish to retain in the form of money in different sets of circumstances?

We shall find that the mistake in the accepted theories of the rate of interest lies in their attempting to derive the rate of interest from the first of these two constituents of psychological time-preference to the neglect of the second; and it is this neglect which we must endeavour to repair.

It should be obvious that the rate of interest cannot

be a return to saving or waiting as such. For if a man hoards his savings in cash, he earns no interest, though he saves just as much as before. On the contrary, the mere definition of the rate of interest tells us in so many words that the rate of interest is the reward for parting with liquidity for a specified period. For the rate of interest is, in itself, nothing more than the inverse proportion between a sum of money and what can be obtained for parting with control over the money in exchange for a debt[1] for a stated period of time.[2]

Thus the rate of interest at any time, being the reward for parting with liquidity, is a measure of the unwillingness of those who possess money to part with their liquid control over it. The rate of interest is not the 'price' which brings into equilibrium the demand for resources to invest with the readiness to abstain from present consumption. It is the 'price' which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash;—which implies that if the rate of interest were lower, i.e. if the reward for parting with cash were diminished, the aggregate amount of cash which the public would wish to hold would exceed the available supply, and that if the rate of interest were raised, there would be a surplus of cash which no one would be willing to hold. If this explanation is correct, the quantity of money is the

[1] Without disturbance to this definition, we can draw the line between 'money' and 'debts' at whatever point is most convenient for handling a particular problem. For example, we can treat as *money* any command over general purchasing power which the owner has not parted with for a period in excess of three months, and as *debt* what cannot be recovered for a longer period than this; or we can substitute for 'three months' one month or three days or three hours or any other period; or we can exclude from *money* whatever is not legal tender on the spot. It is often convenient in practice to include in *money* time-deposits with banks and, occasionally, even such instruments as (e.g.) treasury bills. As a rule, I shall, as in my *Treatise on Money*, assume that money is co-extensive with bank deposits.

[2] In general discussion, as distinct from specific problems where the period of the debt is expressly specified, it is convenient to mean by the rate of interest the complex of the various rates of interest current for different periods of time, i.e. for debts of different maturities.

other factor, which, in conjunction with liquidity-preference, determines the actual rate of interest in given circumstances. Liquidity-preference is a potentiality or functional tendency, which fixes the quantity of money which the public will hold when the rate of interest is given; so that if r is the rate of interest, M the quantity of money and L the function of liquidity-preference, we have $M = L(r)$. This is where, and how, the quantity of money enters into the economic scheme.

At this point, however, let us turn back and consider why such a thing as liquidity-preference exists. In this connection we can usefully employ the ancient distinction between the use of money for the transaction of current business and its use as a store of wealth. As regards the first of these two uses, it is obvious that up to a point it is worth while to sacrifice a certain amount of interest for the convenience of liquidity. But, given that the rate of interest is never negative, why should anyone prefer to hold his wealth in a form which yields little or no interest to holding it in a form which yields interest (assuming, of course, at this stage, that the risk of default is the same in respect of a bank balance as of a bond)? A full explanation is complex and must wait for chapter 15. There is, however, a necessary condition failing which the existence of a liquidity-preference for money as a means of holding wealth could not exist.

This necessary condition is the existence of *uncertainty* as to the future of the rate of interest, i.e. as to the complex of rates of interest for varying maturities which will rule at future dates. For if the rates of interest ruling at all future times could be foreseen with certainty, all future rates of interest could be inferred from the *present* rates of interest for debts of different maturities, which would be adjusted to the knowledge of the future rates. For example, if ${}_1d_r$ is the value in the present year 1 of £1 deferred r years and it is

known that ${}_nd_r$ will be the value in the year n of £1 deferred r years from that date, we have

$$ {}_nd_r = \frac{{}_1d_{n+r}}{{}_1d_n}; $$

whence it follows that the rate at which any debt can be turned into cash n years hence is given by two out of the complex of current rates of interest. If the current rate of interest is positive for debts of every maturity, it must always be more advantageous to purchase a debt than to hold cash as a store of wealth.

If, on the contrary, the future rate of interest is uncertain we cannot safely infer that ${}_nd_r$ will prove to be equal to $\frac{{}_1d_{n+r}}{{}_1d_n}$ when the time comes. Thus if a need for liquid cash may conceivably arise before the expiry of n years, there is a risk of a loss being incurred in purchasing a long-term debt and subsequently turning it into cash, as compared with holding cash. The actuarial profit or mathematical expectation of gain calculated in accordance with the existing probabilities—if it can be so calculated, which is doubtful—must be sufficient to compensate for the risk of disappointment.

There is, moreover, a further ground for liquidity-preference which results from the existence of uncertainty as to the future of the rate of interest, provided that there is an organised market for dealing in debts. For different people will estimate the prospects differently and anyone who differs from the predominant opinion as expressed in market quotations may have a good reason for keeping liquid resources in order to profit, if he is right, from its turning out in due course that the ${}_1d_r$'s were in a mistaken relationship to one another.[1]

This is closely analogous to what we have already

[1] This is the same point as I discussed in my *Treatise on Money* under the designation of the two views and the 'bull-bear' position.

discussed at some length in connection with the marginal efficiency of capital. Just as we found that the marginal efficiency of capital is fixed, not by the 'best' opinion, but by the market valuation as determined by mass psychology, so also expectations as to the future of the rate of interest as fixed by mass psychology have their reactions on liquidity-preference;—but with this addition that the individual, who believes that future rates of interest will be above the rates assumed by the market, has a reason for keeping actual liquid cash,[1] whilst the individual who differs from the market in the other direction will have a motive for borrowing money for short periods in order to purchase debts of longer term. The market price will be fixed at the point at which the sales of the 'bears' and the purchases of the 'bulls' are balanced.

The three divisions of liquidity-preference which we have distinguished above may be defined as depending on (i) the transactions-motive, i.e. the need of cash for the current transaction of personal and business exchanges; (ii) the precautionary-motive, i.e. the desire for security as to the future cash equivalent of a certain proportion of total resources; and (iii) the speculative-motive, i.e. the object of securing profit from knowing better than the market what the future will bring forth. As when we were discussing the marginal efficiency of capital, the question of the desirability of having a highly organised market for dealing with debts presents us with a dilemma. For, in the absence of an organised market, liquidity-preference due to the precautionary-motive would be greatly increased; whereas the existence of an organised market gives an

[1] It might be thought that, in the same way, an individual, who believed that the prospective yield of investments will be below what the market is expecting, will have a sufficient reason for holding liquid cash. But this is not the case. He has a sufficient reason for holding cash or debts in preference to equities; but the purchase of debits will be a preferable alternative to holding cash, unless he also believes that the future rate of interest will prove to be higher than the market is supposing.

opportunity for wide fluctuations in liquidity-preference due to the speculative-motive.

It may illustrate the argument to point out that, if the liquidity-preferences due to the transactions-motive and the precautionary-motive are assumed to absorb a quantity of cash which is not very sensitive to changes in the rate of interest as such and apart from its reactions on the level of income, so that the total quantity of money, less this quantity, is available for satisfying liquidity-preferences due to the speculative-motive, the rate of interest and the price of bonds have to be fixed at the level at which the desire on the part of certain individuals to hold cash (because at that level they feel 'bearish' of the future of bonds) is exactly equal to the amount of cash available for the speculative-motive. Thus each increase in the quantity of money must raise the price of bonds sufficiently to exceed the expectations of some 'bull' and so influence him to sell his bond for cash and join the 'bear' brigade. If, however, there is a negligible demand for cash from the speculative-motive except for a short transitional interval, an increase in the quantity of money will have to lower the rate of interest almost forthwith, in whatever degree is necessary to raise employment and the wage-unit sufficiently to cause the additional cash to be absorbed by the transactions-motive and the precautionary-motive.

As a rule, we can suppose that the schedule of liquidity-preference relating the quantity of money to the rate of interest is given by a smooth curve which shows the rate of interest falling as the quantity of money is increased. For there are several different causes all leading towards this result.

In the first place, as the rate of interest falls, it is likely, *cet. par.*, that more money will be absorbed by liquidity-preferences due to the transactions-motive. For if the fall in the rate of interest increases the national income, the amount of money which it is convenient to

keep for transactions will be increased more or less proportionately to the increase in income; whilst, at the same time, the cost of the convenience of plenty of ready cash in terms of loss of interest will be diminished. Unless we measure liquidity-preference in terms of wage-units rather than of money (which is convenient in some contexts), similar results follow if the increased employment ensuing on a fall in the rate of interest leads to an increase of wages, i.e. to an increase in the money value of the wage-unit. In the second place, every fall in the rate of interest may, as we have just seen, increase the quantity of cash which certain individuals will wish to hold because their views as to the future of the rate of interest differ from the market views.

Nevertheless, circumstances can develop in which even a large increase in the quantity of money may exert a comparatively small influence on the rate of interest. For a large increase in the quantity of money may cause so much uncertainty about the future that liquidity-preferences due to the precautionary-motive may be strengthened; whilst opinion about the future of the rate of interest may be so unanimous that a small change in present rates may cause a mass movement into cash. It is interesting that the stability of the system and its sensitiveness to changes in the quantity of money should be so dependent on the existence of a *variety* of opinion about what is uncertain. Best of all that we should know the future. But if not, then, if we are to control the activity of the economic system by changing the quantity of money, it is important that opinions should differ. Thus this method of control is more precarious in the United States, where everyone tends to hold the same opinion at the same time, than in England where differences of opinion are more usual.

III

We have now introduced money into our causal nexus for the first time, and we are able to catch a first glimpse of the way in which changes in the quantity of money work their way into the economic system. If, however, we are tempted to assert that money is the drink which stimulates the system to activity, we must remind ourselves that there may be several slips between the cup and the lip. For whilst an increase in the quantity of money may be expected, *cet. par.*, to reduce the rate of interest, this will not happen if the liquidity-preferences of the public are increasing more than the quantity of money; and whilst a decline in the rate of interest may be expected, *cet. par.*, to increase the volume of investment, this will not happen if the schedule of the marginal efficiency of capital is falling more rapidly than the rate of interest; and whilst an increase in the volume of investment may be expected, *cet. par.*, to increase employment, this may not happen if the propensity to consume is falling off. Finally, if employment increases, prices will rise in a degree partly governed by the shapes of the physical supply functions, and partly by the liability of the wage-unit to rise in terms of money. And when output has increased and prices have risen, the effect of this on liquidity-preference will be to increase the quantity of money necessary to maintain a given rate of interest.

IV

Whilst liquidity-preference due to the speculative-motive corresponds to what in my *Treatise on Money* I called 'the state of bearishness', it is by no means the same thing. For 'bearishness' is there defined as the functional relationship, not between the rate of interest (or price of debts) and the quantity of money, but between the price of assets and debts, taken together,

money, but may also give rise to changed expectations concerning the future policy of the central bank or of the government. Changes in the liquidity function itself, due to a change in the news which causes revision of expectations, will often be discontinuous, and will, therefore, give rise to a corresponding discontinuity of change in the rate of interest. Only, indeed, in so far as the change in the news is differently interpreted by different individuals or affects individual interests differently will there be room for any increased activity of dealing in the bond market. If the change in the news affects the judgment and the requirements of everyone in precisely the same way, the rate of interest (as indicated by the prices of bonds and debts) will be adjusted forthwith to the new situation without any market transactions being necessary.

Thus, in the simplest case, where everyone is similar and similarly placed, a change in circumstances or expectations will not be capable of causing any displacement of money whatever;—it will simply change the rate of interest in whatever degree is necessary to offset the desire of each individual, felt at the previous rate, to change his holding of cash in response to the new circumstances or expectations; and, since everyone will change his ideas as to the rate which would induce him to alter his holdings of cash in the same degree, no transactions will result. To each set of circumstances and expectations there will correspond an appropriate rate of interest, and there will never be any question of anyone changing his usual holdings of cash.

In general, however, a change in circumstances or expectations will cause some realignment in individual holdings of money;—since, in fact, a change will influence the ideas of different individuals differently by reasons partly of differences in environment and the reason for which money is held and partly of differences in knowledge and interpretation of the

new situation. Thus the new equilibrium rate of interest will be associated with a redistribution of money-holdings. Nevertheless it is the change in the rate of interest, rather than the redistribution of cash, which deserves our main attention. The latter is incidental to individual differences, whereas the essential phenomenon is that which occurs in the simplest case. Moreover, even in the general case, the shift in the rate of interest is usually the most prominent part of the reaction to a change in the news. The movement in bond-prices is, as the newspapers are accustomed to say, 'out of all proportion to the activity of dealing';—which is as it should be, in view of individuals being much more similar than they are dissimilar in their reaction to news.

II

Whilst the amount of cash which an individual decides to hold to satisfy the transactions-motive and the precautionary-motive is not entirely independent of what he is holding to satisfy the speculative-motive, it is a safe first approximation to regard the amounts of these two sets of cash-holdings as being largely independent of one another. Let us, therefore, for the purposes of our further analysis, break up our problem in this way.

Let the amount of cash held to satisfy the transactions- and precautionary-motives be M_1, and the amount held to satisfy the speculative-motive be M_2. Corresponding to these two compartments of cash, we then have two liquidity functions L_1 and L_2. L_1 mainly depends on the level of income, whilst L_2 mainly depends on the relation between the current rate of interest and the state of expectation. Thus

$$M = M_1 + M_2 = L_1(Y) + L_2(r),$$

where L_1 is the liquidity function corresponding to

an income Y, which determines M_1, and L_2 is the liquidity function of the rate of interest r, which determines M_2. It follows that there are three matters to investigate: (i) the relation of changes in M to Y and r, (ii) what determines the shape of L_1, (iii) what determines the shape of L_2.

(i) The relation of changes in M to Y and r depends, in the first instance, on the way in which changes in M come about. Suppose that M consists of gold coins and that changes in M can only result from increased returns to the activities of gold-miners who belong to the economic system under examination. In this case changes in M are, in the first instance, directly associated with changes in Y, since the new gold accrues as someone's income. Exactly the same conditions hold if changes in M are due to the government printing money wherewith to meet its current expenditure;—in this case also the new money accrues as someone's income. The new level of income, however, will not continue sufficiently high for the requirements of M_1 to absorb the whole of the increase in M; and some portion of the money will seek an outlet in buying securities or other assets until r has fallen so as to bring about an increase in the magnitude of M_2 and at the same time to stimulate a rise in Y to such an extent that the new money is absorbed either in M_2 or in the M_1 which corresponds to the rise in Y caused by the fall in r. Thus at one remove this case comes to the same thing as the alternative case, where the new money can only be issued in the first instance by a relaxation of the conditions of credit by the banking system, so as to induce someone to sell the banks a debt or a bond in exchange for the new cash.

It will, therefore, be safe for us to take the latter case as typical. A change in M can be assumed to operate by changing r, and a change in r will lead to a new equilibrium partly by changing M_2 and partly

by changing Y and therefore M_1. The division of the increment of cash between M_1 and M_2 in the new position of equilibrium will depend on the responses of investment to a reduction in the rate of interest and of income to an increase in investment.[1] Since Y partly depends on r, it follows that a given change in M has to cause a sufficient change in r for the resultant changes in M_1 and M_2 respectively to add up to the given change in M.

(ii) It is not always made clear whether the income-velocity of money is defined as the ratio of Y to M or as the ratio of Y to M_1. I propose, however, to take it in the latter sense. Thus if V is the income-velocity of money,

$$L_1(Y) = \frac{Y}{V} = M_1.$$

There is, of course, no reason for supposing that V is constant. Its value will depend on the character of banking and industrial organisation, on social habits, on the distribution of income between different classes and on the effective cost of holding idle cash. Nevertheless, if we have a short period of time in view and can safely assume no material change in any of these factors, we can treat V as nearly enough constant.

(iii) Finally there is the question of the relation between M_2 and r. We have seen in chapter 13 that *uncertainty* as to the future course of the rate of interest is the sole intelligible explanation of the type of liquidity-preference L_2 which leads to the holding of cash M_2. It follows that a given M_2 will not have a definite quantitative relation to a given rate of interest of r;—what matters is not the *absolute* level of r but the degree of its divergence from what is considered a fairly *safe* level of r, having regard to those calculations of probability which are being relied on. Nevertheless, there are two reasons for expecting that, in

[1] We must postpone to Book V the question of what will determine the character of the new equilibrium.

any given state of expectation, a fall in r will be associated with an increase in M_2. In the first place, if the general view as to what is a safe level of r is unchanged, every fall in r reduces the market rate relatively to the 'safe' rate and therefore increases the risk of illiquidity; and, in the second place, every fall in r reduces the current earnings from illiquidity, which are available as a sort of insurance premium to offset the risk of loss on capital account, by an amount equal to the difference between the *squares* of the old rate of interest and the new. For example, if the rate of interest on a long-term debt is 4 per cent, it is preferable to sacrifice liquidity unless on a balance of probabilities it is feared that the long-term rate of interest may rise faster than by 4 per cent of itself per annum, i.e. by an amount greater than 0·16 per cent per annum. If, however, the rate of interest is already as low as 2 per cent, the running yield will only offset a rise in it of as little as 0·04 per cent per annum. This, indeed, is perhaps the chief obstacle to a fall in the rate of interest to a very low level. Unless reasons are believed to exist why future experience will be very different from past experience, a long-term rate of interest of (say) 2 per cent leaves more to fear than to hope, and offers, at the same time, a running yield which is only sufficient to offset a very small measure of fear.

It is evident, then, that the rate of interest is a highly psychological phenomenon. We shall find, indeed, in Book V that it cannot be in equilibrium at a level *below* the rate which corresponds to full employment; because at such a level a state of true inflation will be produced, with the result that M_1 will absorb ever-increasing quantities of cash. But at a level *above* the rate which corresponds to full employment, the long-term market-rate of interest will depend, not only on the current policy of the monetary authority, but also on market expectations concerning its future policy. The

short-term rate of interest is easily controlled by the monetary authority, both because it is not difficult to produce a conviction that its policy will not greatly change in the very near future, and also because the possible loss is small compared with the running yield (unless it is approaching vanishing point). But the long-term rate may be more recalcitrant when once it has fallen to a level which, on the basis of past experience and present expectations of *future* monetary policy, is considered 'unsafe' by representative opinion. For example, in a country linked to an international gold standard, a rate of interest lower than prevails elsewhere will be viewed with a justifiable lack of confidence; yet a domestic rate of interest dragged up to a parity with the *highest* rate (highest after allowing for risk) prevailing in any country belonging to the international system may be much higher than is consistent with domestic full employment.

Thus a monetary policy which strikes public opinion as being experimental in character or easily liable to change may fail in its objective of greatly reducing the long-term rate of interest, because M_2 may tend to increase almost without limit in response to a reduction of r below a certain figure. The same policy, on the other hand, may prove easily successful if it appeals to public opinion as being reasonable and practicable and in the public interest, rooted in strong conviction, and promoted by an authority unlikely to be superseded.

It might be more accurate, perhaps, to say that the rate of interest is a highly conventional, rather than a highly psychological, phenomenon. For its actual value is largely governed by the prevailing view as to what its value is expected to be. *Any* level of interest which is accepted with sufficient conviction as *likely* to be durable *will* be durable; subject, of course, in a changing society to fluctuations for all kinds of reasons round the expected normal. In particular, when M_1

is increasing faster than M, the rate of interest will rise, and *vice versa*. But it may fluctuate for decades about a level which is chronically too high for full employment;—particularly if it is the prevailing opinion that the rate of interest is self-adjusting, so that the level established by convention is thought to be rooted in objective grounds much stronger than convention, the failure of employment to attain an optimum level being in no way associated, in the minds either of the public or of authority, with the prevalence of an inappropriate range of rates of interest.

The difficulties in the way of maintaining effective demand at a level high enough to provide full employment, which ensue from the association of a conventional and fairly stable long-term rate of interest with a fickle and highly unstable marginal efficiency of capital, should be, by now, obvious to the reader.

Such comfort as we can fairly take from more encouraging reflections must be drawn from the hope that, precisely because the convention is not rooted in secure knowledge, it will not be always unduly resistant to a modest measure of persistence and consistency of purpose by the monetary authority. Public opinion can be fairly rapidly accustomed to a modest fall in the rate of interest and the conventional expectation of the future may be modified accordingly; thus preparing the way for a further movement—up to a point. The fall in the long-term rate of interest in Great Britain after her departure from the gold standard provides an interesting example of this;—the major movements were effected by a series of discontinuous jumps, as the liquidity function of the public, having become accustomed to each successive reduction, became ready to respond to some new incentive in the news or in the policy of the authorities.

III

We can sum up the above in the proposition that in any given state of expectation there is in the minds of the public a certain potentiality towards holding cash beyond what is required by the transactions-motive or the precautionary-motive, which will realise itself in actual cash-holdings in a degree which depends on the terms on which the monetary authority is willing to create cash. It is this potentiality which is summed up in the liquidity function L_2.

Corresponding to the quantity of money created by the monetary authority, there will, therefore, be *cet. par.* a determinate rate of interest or, more strictly, a determinate complex of rates of interest for debts of different maturities. The same thing, however, would be true of any other factor in the economic system taken separately. Thus this particular analysis will only be useful and significant in so far as there is some specially direct or purposive connection between changes in the quantity of money and changes in the rate of interest. Our reason for supposing that there is such a special connection arises from the fact that, broadly speaking, the banking system and the monetary authority are dealers in money and debts and not in assets or consumables.

If the monetary authority were prepared to deal both ways on specified terms in debts of all maturities, and even more so if it were prepared to deal in debts of varying degrees of risk, the relationship between the complex of rates of interest and the quantity of money would be direct. The complex of rates of interest would simply be an expression of the terms on which the banking system is prepared to acquire or part with debts; and the quantity of money would be the amount which can find a home in the possession of individuals who—after taking account of all relevant circumstances—prefer the control of liquid cash to parting with it

Chapter 22

NOTES ON THE TRADE CYCLE

Since we claim to have shown in the preceding chapters what determines the volume of employment at any time, it follows, if we are right, that our theory must be capable of explaining the phenomena of the trade cycle.

If we examine the details of any actual instance of the trade cycle, we shall find that it is highly complex and that every element in our analysis will be required for its complete explanation. In particular we shall find that fluctuations in the propensity to consume, in the state of liquidity-preference, and in the marginal efficiency of capital have all played a part. But I suggest that the essential character of the trade cycle and, especially, the regularity of time-sequence and of duration which justifies us in calling it a *cycle*, is mainly due to the way in which the marginal efficiency of capital fluctuates. The trade cycle is best regarded, I think, as being occasioned by a cyclical change in the marginal efficiency of capital, though complicated and often aggravated by associated changes in the other significant short-period variables of the economic system. To develop this thesis would occupy a book rather than a chapter, and would require a close examination of facts. But the following short notes will be sufficient to indicate the line of investigation which our preceding theory suggests.

I

By a *cyclical* movement we mean that as the system progresses in, e.g. the upward direction, the forces

propelling it upwards at first gather force and have a cumulative effect on one another but gradually lose their strength until at a certain point they tend to be replaced by forces operating in the opposite direction; which in turn gather force for a time and accentuate one another, until they too, having reached their maximum development, wane and give place to their opposite. We do not, however, merely mean by a *cyclical* movement that upward and downward tendencies, once started, do not persist for ever in the same direction but are ultimately reversed. We mean also that there is some recognisable degree of regularity in the time-sequence and duration of the upward and downward movements.

There is, however, another characteristic of what we call the trade cycle which our explanation must cover if it is to be adequate; namely, the phenomenon of the *crisis*—the fact that the substitution of a downward for an upward tendency often takes place suddenly and violently, whereas there is, as a rule, no such sharp turning-point when an upward is substituted for a downward tendency.

Any fluctuation in investment not offset by a corresponding change in the propensity to consume will, of course, result in a fluctuation in employment. Since, therefore, the volume of investment is subject to highly complex influences, it is highly improbable that all fluctuations either in investment itself or in the marginal efficiency of capital will be of a cyclical character. One special case, in particular, namely, that which is associated with agricultural fluctuations, will be separately considered in a later section of this chapter. I suggest, however, that there are certain definite reasons why, in the case of a typical industrial trade cycle in the nineteenth-century environment, fluctuations in the marginal efficiency of capital should have had cyclical characteristics. These reasons are by no means unfamiliar either in themselves or as explanations of the trade

cycle. My only purpose here is to link them up with the preceding theory.

II

I can best introduce what I have to say by beginning with the later stages of the boom and the onset of the 'crisis'.

We have seen above that the marginal efficiency of capital[1] depends, not only on the existing abundance or scarcity of capital-goods and the current cost of production of capital-goods, but also on current expectations as to the future yield of capital-goods. In the case of durable assets it is, therefore, natural and reasonable that expectations of the future should play a dominant part in determining the scale on which new investment is deemed advisable. But, as we have seen, the basis for such expectations is very precarious. Being based on shifting and unreliable evidence, they are subject to sudden and violent changes.

Now, we have been accustomed in explaining the 'crisis' to lay stress on the rising tendency of the rate of interest under the influence of the increased demand for money both for trade and speculative purposes. At times this factor may certainly play an aggravating and, occasionally perhaps, an initiating part. But I suggest that a more typical, and often the predominant, explanation of the crisis is, not primarily a rise in the rate of interest, but a sudden collapse in the marginal efficiency of capital.

The later stages of the boom are characterised by optimistic expectations as to the future yield of capital-goods sufficiently strong to offset their growing abundance and their rising costs of production and, probably, a rise in the rate of interest also. It is of the nature of

[1] It is often convenient in contexts where there is no room for misunderstanding to write 'the marginal efficiency of capital', where 'the schedule of the marginal efficiency of capital' is meant.

organised investment markets, under the influence of purchasers largely ignorant of what they are buying and of speculators who are more concerned with forecasting the next shift of market sentiment than with a reasonable estimate of the future yield of capital-assets, that, when disillusion falls upon an over-optimistic and over-bought market, it should fall with sudden and even catastrophic force.[1] Moreover, the dismay and uncertainty as to the future which accompanies a collapse in the marginal efficiency of capital naturally precipitates a sharp increase in liquidity-preference—and hence a rise in the rate of interest. Thus the fact that a collapse in the marginal efficiency of capital tends to be associated with a rise in the rate of interest may seriously aggravate the decline in investment. But the essence of the situation is to be found, nevertheless, in the collapse in the marginal efficiency of capital, particularly in the case of those types of capital which have been contributing most to the previous phase of heavy new investment. Liquidity-preference, except those manifestations of it which are associated with increasing trade and speculation, does not increase until *after* the collapse in the marginal efficiency of capital.

It is this, indeed, which renders the slump so intractable. Later on, a decline in the rate of interest will be a great aid to recovery and, probably, a necessary condition of it. But, for the moment, the collapse in the marginal efficiency of capital may be so complete that no practicable reduction in the rate of interest will be enough. If a reduction in the rate of interest was capable of proving an effective remedy by itself, it might be possible to achieve a recovery without the elapse of any considerable interval of time and by means more or less directly under the control of the monetary

[1] I have shown above (chapter 12) that, although the private investor is seldom himself directly responsible for new investment, nevertheless the entrepreneurs, who are directly responsible, will find it financially advantageous, and often unavoidable, to fall in with the ideas of the market, even though they themselves are better instructed.

authority. But, in fact, this is not usually the case; and it is not so easy to revive the marginal efficiency of capital, determined, as it is, by the uncontrollable and disobedient psychology of the business world. It is the return of confidence, to speak in ordinary language, which is so insusceptible to control in an economy of individualistic capitalism. This is the aspect of the slump which bankers and business men have been right in emphasising, and which the economists who have put their faith in a 'purely monetary' remedy have underestimated.

This brings me to my point. The explanation of the *time-element* in the trade cycle, of the fact that an interval of time of a particular order of magnitude must usually elapse before recovery begins, is to be sought in the influences which govern the recovery of the marginal efficiency of capital. There are reasons, given firstly by the length of life of durable assets in relation to the normal rate of growth in a given epoch, and secondly by the carrying-costs of surplus stocks, why the duration of the downward movement should have an order of magnitude which is not fortuitous, which does not fluctuate between, say, one year this time and ten years next time, but which shows some regularity of habit between, let us say, three and five years.

Let us recur to what happens at the crisis. So long as the boom was continuing, much of the new investment showed a not unsatisfactory current yield. The disillusion comes because doubts suddenly arise concerning the reliability of the prospective yield, perhaps because the current yield shows signs of falling off, as the stock of newly produced durable goods steadily increases. If current costs of production are thought to be higher than they will be later on, that will be a further reason for a fall in the marginal efficiency of capital. Once doubt begins it spreads rapidly. Thus at the outset of the slump there is probably much capital of which the marginal efficiency has become negligible or even

negative. But the interval of time, which will have to elapse before the shortage of capital through use, decay and obsolescence causes a sufficiently obvious scarcity to increase the marginal efficiency, may be a somewhat stable function of the average durability of capital in a given epoch. If the characteristics of the epoch shift, the standard time-interval will change. If, for example, we pass from a period of increasing population into one of declining population, the characteristic phase of the cycle will be lengthened. But we have in the above a substantial reason why the duration of the slump should have a definite relationship to the length of life of durable assets and to the normal rate of growth in a given epoch.

The second stable time-factor is due to the carrying-costs of surplus stocks which force their absorption within a certain period, neither very short nor very long. The sudden cessation of new investment after the crisis will probably lead to an accumulation of surplus stocks of unfinished goods. The carrying-costs of these stocks will seldom be less than 10 per cent. per annum. Thus the fall in their price needs to be sufficient to bring about a restriction which provides for their absorption within a period of, say, three to five years at the outside. Now the process of absorbing the stocks represents negative investment, which is a further deterrent to employment; and, when it is over, a manifest relief will be experienced.

Moreover, the reduction in working capital, which is necessarily attendant on the decline in output on the downward phase, represents a further element of disinvestment, which may be large; and, once the recession has begun, this exerts a strong cumulative influence in the downward direction. In the earliest phase of a typical slump there will probably be an investment in increasing stocks which helps to offset disinvestment in working-capital; in the next phase there may be a short period of disinvestment both in stocks and in working-

capital; after the lowest point has been passed there is likely to be a further disinvestment in stocks which partially offsets reinvestment in working-capital; and, finally, after the recovery is well on its way, both factors will be simultaneously favourable to investment. It is against this background that the additional and superimposed effects of fluctuations of investment in durable goods must be examined. When a decline in this type of investment has set a cyclical fluctuation in motion there will be little encouragement to a recovery in such investment until the cycle has partly run its course.[1]

Unfortunately a serious fall in the marginal efficiency of capital also tends to affect adversely the propensity to consume. For it involves a severe decline in the market value of stock exchange equities. Now, on the class who take an active interest in their stock exchange investments, especially if they are employing borrowed funds, this naturally exerts a very depressing influence. These people are, perhaps, even more influenced in their readiness to spend by rises and falls in the value of their investments than by the state of their incomes. With a 'stock-minded' public as in the United States to-day, a rising stock-market may be an almost essential condition of a satisfactory propensity to consume; and this circumstance, generally overlooked until lately, obviously serves to aggravate still further the depressing effect of a decline in the marginal efficiency of capital.

When once the recovery has been started, the manner in which it feeds on itself and cumulates is obvious. But during the downward phase, when both fixed capital and stocks of materials are for the time being redundant and working-capital is being reduced, the schedule of the marginal efficiency of capital may fall so low that it can scarcely be corrected, so as to

[1] Some part of the discussion in my *Treatise on Money* [*JMK*, vol. v], Book IV, bears upon the above.

secure a satisfactory rate of new investment, by any practicable reduction in the rate of interest. Thus with markets organised and influenced as they are at present, the market estimation of the marginal efficiency of capital may suffer such enormously wide fluctuations that it cannot be sufficiently offset by corresponding fluctuations in the rate of interest. Moreover, the corresponding movements in the stock-market may, as we have seen above, depress the propensity to consume just when it is most needed. In conditions of *laissez-faire* the avoidance of wide fluctuations in employment may, therefore, prove impossible without a far-reaching change in the psychology of investment markets such as there is no reason to expect. I conclude that the duty of ordering the current volume of investment cannot safely be left in private hands.

III

The preceding analysis may appear to be in conformity with the view of those who hold that over-investment is the characteristic of the boom, that the avoidance of this over-investment is the only possible remedy for the ensuing slump, and that, whilst for the reasons given above the slump cannot be prevented by a low rate of interest, nevertheless the boom can be avoided by a high rate of interest. There is, indeed, force in the argument that a high rate of interest is much more effective against a boom than a low rate of interest against a slump.

To infer these conclusions from the above would, however, misinterpret my analysis; and would, according to my way of thinking, involve serious error. For the term over-investment is ambiguous. It may refer to investments which are destined to disappoint the expectations which prompted them or for which there is no use in conditions of severe unemployment, or it may indicate a state of affairs where every kind of capital-

goods is so abundant that there is no new investment which is expected, even in conditions of full employment, to earn in the course of its life more than its replacement cost. It is only the latter state of affairs which is one of over-investment, strictly speaking, in the sense that any further investment would be a sheer waste of resources.[1] Moreover, even if over-investment in this sense was a normal characteristic of the boom, the remedy would not lie in clapping on a high rate of interest which would probably deter some useful investments and might further diminish the propensity to consume, but in taking drastic steps, by redistributing incomes or otherwise, to stimulate the propensity to consume.

According to my analysis, however, it is only in the former sense that the boom can be said to be characterised by over-investment. The situation, which I am indicating as typical, is not one in which capital is so abundant that the community as a whole has no reasonable use for any more, but where investment is being made in conditions which are unstable and cannot endure, because it is prompted by expectations which are destined to disappointment.

It may, of course, be the case—indeed it is likely to be—that the illusions of the boom cause particular types of capital-assets to be produced in such excessive abundance that some part of the output is, on any criterion, a waste of resources;—which sometimes happens, we may add, even when there is no boom. It leads, that is to say, to *misdirected* investment. But over and above this it is an essential characteristic of the boom that investments which will in fact yield, say, 2 per cent in conditions of full employment are made in the expectation of a yield of, say, 6 per cent, and are valued accordingly. When the disillusion comes, this expectation is

[1] On certain assumptions, however, as to the distribution of the propensity to consume through time, investment which yielded a negative return might be advantageous in the sense that, for the community as a whole, it would maximise satisfaction.