John Maynard Keynes

The General Theory of Employment, Interest, and Money

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. 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 employment 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 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.

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. For entrepreneurs will endeavour to fix the amount of employment 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 $Z = \varphi(N)$, which can be called the *aggregate supply function*. 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 $f(\mathrm{N})$ and $\varphi(\mathrm{N})$ are equal for *all* values of N, i.e. for all levels of output and employment; and that when there is an increase in Z $(= \varphi(N))$ 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 consurnption 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 $D_1 + D_2 = D = \varphi(N)$, where 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 $\varphi(N) - \chi(N) = D_2$.

(5) Hence the volume of employment in equilibrium depends on (i) the aggregate supply function, (ii) the propensity to consume, 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 $D = \varphi(N)$ 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_1will 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 employment 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 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, 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 manner of the classical postulates. It may well be that the classical theory represents the way in which we should like our economy to behave. But to assume that it actually does so is to assume our difficulties away.

Rendered into HTML on Sun Apr 20 09:55:24 2003, by Steve Thomas for The University of Adelaide Library Electronic Texts Collection.