10

THE RELATION OF ECONOMIC GROWTH AND CYCLICAL FLUCTUATIONS[1]

SINCE the very beginnings of speculation on the problem of the trade cycle, the cyclical swings of the economic system have been regarded as being inherently connected with the essentially "dynamic" process of economic growth. For this there were two reasons. In the first place, with relatively unimportant exceptions (such as Jevons' sun spots), even the early writers on the trade cycle (like Marx, Aftalion, Spiethoff or J. A. Hobson) emphasised the key rôle of the investment process, and of the growth of productive capacity, in the generation of cyclical movements. Secondly, the close connection between the cycle and dynamic change was "visible to the naked eye"—inasmuch as the historical boom periods of the last 150 years were, in many cases, clearly associated with the exploitation of major technical innovations, such as the railways, electricity or the automobile.

At least one distinguished author—Joseph Schumpeter—put forward a trade-cycle theory which makes the cycle itself simply a by-product of economic progress—booms and depressions being "the form which progress takes in a capitalist society". This is because, according to Schumpeter, the realisation of major innovations must await a time when the general economic climate is favourable to them; the adoption of major innovations by the heroic innovating entrepreneurs is followed by a host of imitators, giving rise to an investment boom; when the innovations are thus fully exploited (or over-exploited) the economy once more relapses into a depression, until the accumulation of new ideas creates the favourable climate for a new burst of "innovating" investment.

The trouble with Schumpeter's theory is that it is descriptive

[1] Reprinted from the *Economic Journal*, March, 1954. Originally a lecture delivered to the Institut de Science Economique Appliquée in Paris on May 23rd, 1953.

rather than analytical. Although it is easy enough to see how one particular part of the story follows from the preceding part, it is not possible to make the story as a whole into a "model" (meaning by a model the sum total of assumptions which are just sufficient —no more and no less—together to provide the necessary and sufficient conditions for the generation of a recurrent cycle with a clear periodicity) without incorporating into it elements which would suffice by themselves to explain the cycle—without recourse to Schumpeter's own stage army of initiators and imitators, or even the very concept of technical progress. For the necessary "bunching" (in time) of innovating investment, which is essential to Schumpeter's theory, cannot be satisfactorily explained without bringing the Keynesian multiplier, and some variant of the output-investment relationship, to one's aid.

Indeed, the development of trade-cycle theories that followed Keynes' *General Theory* has proved to be positively inimical to the idea that cycle and dynamic growth are inherently connected analytically—to the idea, that is, that the cycle is a mere by-product of, and could not occur in the absence of, "progress". For it has been repeatedly (and in my view, conclusively)[1] shown that a few simple additions to Keynes' own model of a general equilibrium of production in the economy will produce the result that this "equilibrium" will take the form, not of a simple steady rate of production in time, but of a rhythmical movement of constant amplitude and period—in other words, a perpetual oscillation around a stationary equilibrium position. The necessary additions to Keynes' own framework of assumptions to get this result are few and simple. All that is required is to treat the investment demand schedule (Keynes' "marginal efficiency of capital") not, as Keynes did, as an independent variable, but as a function of both the existing stock of capital and of the current rate of output, or simply of the ratio of output to the stock of capital. This is justified, if future profit-expectations can be

[1] Cf. my own paper "A Model of the Trade Cycle", *Economic Journal*, March, 1940, [above pp. 177-92]; which, of course, drew heavily on earlier works of Harrod (*The Trade Cycle*, 1936), Kalecki (*Essays in the Theory of Economic Fluctuations*, 1938) and others. Since that time Richard Goodwin (*Econometrica*, January, 1951, pp. 1-17) and Hicks (*A Contribution to the Theory of the Trade Cycle*, Oxford, 1950) produced cyclical models which in the essential framework of assumptions, though not necessarily in technique, are identical.

regarded as being largely determined by the current rate of profit on capital. For the current rate of profit per unit of capital (or rather per unit of investment) clearly depends on the current relation of output to capital (at any rate if, as a first approximation, profit per unit of output can be taken as given) so that the rate of investment will increase with a rise in output and/or a diminution in the existing stock of capital, and vice versa.[1] Since current output depends (via the Keynesian multiplier) on current investment, whereas the current change in the stock of capital depends on past investment (and will be positive or negative according to whether investment in the previous "period" was greater or less than the necessary replacement of the capital goods currently used up in production), the mechanism will in itself produce a periodic fluctuation of constant amplitude (of the output/capital ratio, of the rate of investment and of the level of output), provided two additional conditions are satisfied:

(1) a change in output (income) is associated with a greater change in the rate of investment than of savings, so that the stability conditions of the "Keynesian" equilibrium[2] are not satisfied for levels of output that imply a rate of investment that is higher than the possible minimum and lower than the possible maximum;[3]

(2) the rate of investment at any one time cannot exceed a certain maximum, or fall below a certain minimum, which

[1] Some writers (notably Harrod and Hicks) prefer to exclude the stock of capital as an explicit variable, and treat instead investment as a function of the *change* in the rate of output (the so-called acceleration principle). It is important to emphasise that the difference is more a matter of methodological and pedagogic convenience than of substance. It is perfectly easy to translate one technique into another without any difference in the results; though the acceleration method requires far more rigid assumptions as regards production techniques and time lags of adjustment (cf. my review article of Hicks' book in the *Economic Journal*, December, 1951), [above pp. 198-202].

[2] The Keynesian stability conditions are simply that a rise in production should increase the aggregate demand for output by less than the value of the additional output (i.e. the "aggregate supply price"). Since the marginal propensity to consume is less than 1 (i.e. the additional demand for consumers' goods generated by a rise in income falls short of the rise in income), this condition will be satisfied if: (*a*) the rate of investment is invariant with respect to changes in output; (*b*) the increase in the rate of investment generated by a rise in output is less than the shortfall in marginal consumers' demand over marginal income—i.e. savings out of marginal income.

[3] Actually it is sufficient for the model if this condition is satisfied in the neighbourhood of the hypothetical "long-run equilibrium" position, where investment=current depreciation. It is simpler to assume, however, that it holds for all levels of investment between a certain minimum and maximum.

implies also that the Keynesian stability conditions must be satisfied (partially at any rate) whenever investment is at the maximum or the minimum level.

The second condition must be satisfied by virtue of the fact that gross investment cannot be less than zero, and cannot be greater than the investment corresponding to full employment, or the capacity of the investment goods industries, whichever is less.[1] The first condition is likely to be satisfied on account of the fact that in situations in which an increase in output induces the creation of additional capacity, the value of that additional capacity is likely to exceed (perhaps several times) the value of the additional output per unit period; hence the additional rate of investment (per unit period) is also likely to exceed the additional savings generated by that additional output (the length of the critical "period" depends on the time taken to adjust current production to changes in demand).[2]

Properties of the "Static" Model

The above framework of assumptions yields a purely "static" model of the trade cycle—"static" in the sense that it accounts

[1] The above statement may be open to criticism on two counts. First, it might be suggested that "full employment" does not set any genuine limit to the rate of investment (in *real* terms), since investment could be increased further at the cost of current consumption, through a process of monetary inflation. Now it is true that by means of inflation it is possible to reduce the proportion of real income consumed. But past experience has repeatedly shown that the extent to which an inflationary process succeeds in augmenting the real rate of savings of the community is very limited; and when, as a result of full employment, the increase in current output slows down, the urge to expand investment further is also reduced. However, there is nothing in the model to prevent an expansionary process from leading to a state of purely monetary inflation (in which case the monetary authorities would be forced to take restrictionist measures sooner or later); the point is rather that it does not make any real difference to the mechanics of the model whether inflation occurs in the course of it or not.

It might also be objected that in the absence of full employment the capacity of the investment goods industries does not set a limit to the expansionary process either: since in that case there would be an urge to undertake investments to expand the capacity of these industries. That is true, but this second type of reaction might operate far too slowly to be of much consequence in relation to the cyclical process. Thus when the demand for new ships exceeds the capacity of the shipbuilding yards the first result is that order books and delivery periods are lengthened. If this lengthening of order books went on for several years the shipbuilding industry would be induced to build new shipyards. It is quite possible, however, that before such major extensions get under way the boom is over.

[2] For a fuller explanation of these conditions cf. my paper in *Economic Journal*, March, 1940, cited above, particularly the Appendix [above pp. 188-92]. As explained there, the advantage of assuming non-linear co-efficients is that it makes it unnecessary to rely on time-lags or to introduce exogenous shocks of any kind in order to account for a cyclical movement of constant amplitude.

for a regular cyclical movement of constant amplitude and period, in the absence of any "dynamic" change, such as technical innovation, population growth or changes in the political or institutional framework, and in which the whole movement, its duration as well as its amplitude, is fully determined by the parameters of the system. It has strictly four phases—an upward phase, a boom phase, a downward phase and a depression phase, the first and third of which may be relatively short in duration in relation to the second and the fourth. Contrary to the general belief (including my own previous views) that under such "static" conditions the duration of the depression phase would necessarily be very much longer than that of the boom phase (owing to the simple fact that capital goods take much longer to wear out than to build), I now believe that the relative duration of the boom phase and the depression phase is simply a matter of the relationship of the output capacity of the investment goods industries to the normal annual depreciation of the capital stock. If this output capacity were twice as large as normal depreciation, the duration of boom and depression phases would be identical; if it were more than twice as large, the depression phase would last longer than the boom phase, while if it were less than twice as large, the depression phase would be shorter than the boom. Hence, without introducing any "dynamic" change, it is possible to construct models in which the depression phase is short relative to the boom phase rather than the other way round. The length of the boom phase depends on a complex set of factors: the annual output capacity of the equipment goods industries, its relation to the capital stock at any one time (which *does* depend both on gestation period and service-life); the "critical" degree of excess capacity which causes entrepreneurs to abstain from adding to capacity further; and finally, the Keynesian multiplier, which determines the extent of utilised capacity at boom and depression levels, respectively. The duration of the depression phase depends on the rate at which existing capacity is reduced by scrapping; and it can be assumed, of course, that it is always the oldest and most obsolete plant which is scrapped. If we suppose that the service-life of equipment produced in any given period is uniquely fixed by technical factors, the equipment

which disappears during any given depression phase is the equipment that was created in the earliest of the boom periods which was still extant at the end of the preceding boom phase; and the period required for its disappearance (assuming both a fixed and a constant service-life for equipment) would depend on both the length of that earlier boom phase and the age-distribution of the capital stock at the end of the previous boom. If the life-time of plant or equipment is not rigidly determined, we cannot take it for granted that the discontinuity in building periods will be reflected in a similar discontinuity of scrapping periods, though in so far as the life-time of equipment can be extended only at a gradually increasing operating cost, there will be an economic motive for keeping obsolete equipment in existence during boom periods and to concentrate the scrapping of equipment on depression periods. Finally, if we suppose that for purely technical reasons, there is a fairly wide spread in the service-life of the equipment created in any given period (both because equipment is of different kinds and also because equipment of any particular kind has a probability-life-distribution rather than a fixed service-life), the disappearance of old equipment will tend to take place at a more or less steady rate in time, irrespective of the discontinuity in building periods; and it may not show much variation as between boom and depression phases. In each case, however, the length of the boom phase in relation to the length of the cycle as a whole will depend simply on the relationship of the average annual depreciation of the capital stock—during the cycle as a whole—to the annual output of capital goods during the boom. If the latter were twice the former, the boom phase would take up one-half of the full cycle-period and so on.

The inter-relation of the various factors could be made clearer in terms of a simple numerical example. Let us suppose that the capital stock consists of ships, while output consists of shipbuilding (investment) and shipping services (consumption); that ships take one year to build and the capacity of the shipbuilding industry is 1 million tons, so that 1 million tons represents also the maximum output of new ships per annum. Let us further suppose that at the boom-level of income (i.e. when the shipbuilding industry is fully utilised), the demand for shipping

services requires a minimum of 12 million tons of capacity and at the depression level of income (when no new ships are built), a minimum of 8 million tons; and finally, that shipowners order new ships whenever the current demand exceeds 80% of their current capacity. On these assumptions the shipping park reaches a maximum of 15 million tons at the end of the boom (125% of 12 million tons) and a minimum of 10 million tons at the end of the depression (125% of 8 million tons): in each case because, at those levels, demand tends to fall below, or rise above, the critical 80% utilisation of capacity. If we now suppose that the service-life of ships is a fixed period of 25 years, the duration of the boom will be 5 years, and the duration of the depression will also be 5 years, since in 5 years the 5 million tons created in the last but three of the building periods (25 years ago) will be scrapped. If we supposed, other things remaining unchanged, the maximum shipbuilding capacity to be 2 million tons, the duration of the boom would be reduced to 2·5 years, but by the same token the duration of the depression would be lengthened to 7·5 years—since, given the size of the shipping park at boom and depression levels, as determined by demand, the age distribution of ships would be such that there would be no ships due for scrapping in the first 5 years of any depression period, and the scrapping would be concentrated in the last 2·5 years of the depression phase, when the 5 million tons of ships built in the boom phase of 25 years before would disappear. Again we could assume that the service-life of ships is 35, 45 or 55 years instead of 25, with similar effects on the relative duration of boom and depression phases; though if the life-time were supposed to be less than 15 years the cyclical mechanism could not go into operation at all unless we also amended some of the other assumptions—e.g. either the minimum demand for shipping services would have to be assumed to be less than the above 8 million tons or the maximum shipbuilding capacity would have to be assumed to be greater than the above 1 million tons per annum.[1]

[1] I am conscious, of course, of the fact that in the above numerical examples the figures were so chosen as to avoid the need for any overlapping as between the building periods and the scrapping periods. Assuming that the service-life of ships is rigidly determined by technical factors, this can be the case only when: (*a*) the boom phase takes up one-half or less of the cycle; and (*b*) the service-life of ships is an integral odd-number multiple of the period of the boom. Thus, if in the above example the

If we now supposed (again, other things remaining unchanged) that while the *average* service-life of ships is 25 years, there is a fairly wide frequency distribution around that average, ships may be assumed to be scrapped at a (more or less) constant rate of 0·5 million tons per annum both during boom and depression phases (corresponding to $\frac{1}{25}$ of the *average* capital stock of 12·5 million tons). In that case the boom phase will last for 10 years (since with an annual output of 1 million tons, it will take 10 years to *raise* the shipping park from 10 to 15 million tons), and the depression phase will also last 10 years (since, with a fixed rate of scrapping of 0·5 million tons, it will take 10 years for the shipping park to be reduced from 15 to 10 million tons). The period of each phase, and of the cycle as a whole, would thus be doubled, as compared with the situation in which (for technical or economic reasons) the actual disappearance of equipment is concentrated on depression phases; but otherwise the mechanism would operate in the same way.[1]

The most important factor in the whole mechanism is the capacity of the capital goods industries which, given the service-life and gestation period of capital goods, determines the rate of capital accumulation during the boom periods. The greater is this factor, the shorter the duration of the boom phase (both absolutely and relatively to the depression phase) and the greater the amplitude of the cyclical movement. The capacity of the capital goods industries relative to the consumer goods industries may in turn be assumed to depend on the proportion of income

service-life of ships were assumed to be a rigid 20 years, there would necessarily be overlapping between building and scrapping periods, which would cause the boom to be prolonged to 15 years, followed by a depression phase of 10 years; if it were to be assumed to be a rigid 30 years, there would be a boom period of 15 years, followed by a depression phase of 20 years; and there may be other more complicated patterns. On the other hand, if the life-time is not assumed to be *narrowly* determined by technical factors—say, in the above example, ships are scrapped when they are over 20 but less than 25 years old—overlapping would be avoided as a result of the operation of economic factors, since there would be an incentive to keep old capacity in existence during the boom, and thus to concentrate scrapping in times of slump.

[1] It is important to emphasise therefore that this particular model is in no way dependent on the so-called "echo-effect"—i.e the bunching of replacement demand in time resulting from past discontinuities in building periods. If there is a wide scatter in the actual service-life of plant and equipment, discontinuities in building periods need not result in corresponding variations in the rate of scrapping—any more than annual fluctuations in the number of births tend to be reflected in corresponding variations in deaths a generation later. But even if the scrapping of equipment is perfectly continuous, there will be a bunching of investment in time as a result of the operation of the multiplier and the "accelerator"

saved at full-employment income levels—in other words, on the short-run savings function. For supposing that maximum investment, as determined by this capacity, were initially less than full-employment savings, there would be an inducement, in each successive boom, to enlarge this capacity, relatively to the rest of the economy, until the deficiency was eliminated. Similarly, if this capacity in relation to the rest of the economy were initially greater than the proportion of income saved at full employment, some part of the capacity of the capital goods industries would become permanently redundant—not being utilised even at the peak of the boom—and would thus tend to disappear. If, however, entrepreneurial expectation in the capital goods industries were more sluggish than in the consumption goods industries, it is quite possible that the capacity of the capital goods industries would be maintained at too low a level to secure full employment in boom periods, and the cyclical movement itself might be weak in amplitude and the booms relatively long in duration.

Looked at in another way, the mechanism depends on the assumption of a short-run savings function which, though it may be represented as being linear, must contain a negative constant. In other words, it is assumed that while the marginal propensity to save is constant, the average propensity to save is a rising function of income in the short period. Hence there is a minimum depression level of income at which (gross) savings are zero and a maximum boom level of savings which is attained when income reaches the full-employment level. If we ignore the fact that, for reasons stated above, the maximum rate of investment during a boom (as determined by the actual capacity of the capital goods industries) may be consistently less than the investment needed for full-employment output, we can simply say that the amplitude and the character of the cyclical movement (i.e. the relative duration of boom and depression phases) is determined simply by the character of the short-run savings function.

Cycle and Trend

The static model outlined above has no trend—which shows that it is not necessary to assume economic growth or dynamic change in order to account for the existence of fluctuations.

Each depression phase lasts precisely as long as is required for the capital stock to fall to the same extent as it had risen in the previous boom; the capital stock, as well as the output, in the corresponding phases of successive cycles are identical. As a pure cyclical model it has therefore little resemblance to the cyclical fluctuations in the real world, the most characteristic feature of which has been that successive booms carried production to successively higher levels; while the creation of capital in boom periods has exceeded many times the net capital depletion in depression periods.

The significant point is, however, that a trend can be incorporated into the above framework of assumptions without upsetting the mechanism; and thus by this simple device it is possible to produce a model whose features show a far better resemblance to the observed phenomena. The model is capable also of absorbing "random shocks" of all kinds, such as wars, revolutions or revolutionary technical changes caused by major inventions, with no more consequence than of displacing (in time) the operation of the cyclical movement. It strongly suggests therefore that economic growth, as well as other dynamic and erratic changes, should be treated rather as the cause of aberrations from the pure rhythm of the static cycle than as the *sine qua non* of the basic phenomena.

The simplest way of introducing a trend is to assume a linear percentage growth in population over time. This will mean that the amount of capital created in each successive boom will be higher than in the previous boom (since in each case the expansion would tend to go on until full employment or full capacity in the investment goods industries is reached; but in this case the capital goods industries would themselves tend to be enlarged in each successive boom), while the capital used up in the succeeding depression phases will be consistently smaller than the capital created in the previous boom phase.[1] It is not certain

[1] To explain this, it is necessary to assume that savings are a function of income per head, rather than total income—i.e. that the savings function also "shifts to the right" with the rise in full-employment output, so that savings always fall to zero, when actual output falls to a certain percentage of full-employment output. Alternatively, it could be assumed (as Hicks assumed) that there is some long-range investment which is autonomous and not responsive to changes in current output, and which is growing at the trend rate.

(because it has not yet been worked out) how the introduction of a logarithmic trend affects the duration of the cycle or the length of the various phases relative to each other. My hunch is that a *logarithmic* trend might leave both these unaffected. While there is more capital created in each succeeding boom, there may also be assumed to be a greater capacity of investment-goods production to create it, so that the length of the boom could remain unchanged; while there is less capital (as a proportion of the total stock) destroyed during the depression, the rate of disappearance of "old" capital is also less (since a greater proportion of the capital stock is of more recent creation), so that the length of the depression phase would also remain unchanged. If this is correct, long booms and short depressions, or vice versa, would be promoted by an accelerating or decelerating trend, rather than by the existence of a trend as such.

It is possible also to make technical progress (in the shape of a linear time-rate of growth in the productivity of labour) the cause of the trend; and so long as this technical progress is assumed to be "neutral"—i.e. leave the ratio of capital and output in production unaffected—its effect on the model is no different from that of population growth. In fact, these two factors are additive—given both neutral technical progress and population growth, the trend will be the resultant of the two, since the rate of increase of full-employment output will be the sum of the rate of increase of productivity and of population. It would also be possible to assume that technical innovation is, on balance, capital using (i.e. involving a net substitution of capital for labour) or else capital saving, though the precise consequences of this on the properties of the model are difficult to work out. Finally, it is possible—as Mr. Goodwin has done in one of his models[1]—to neglect the replacement of capital altogether, and assume, in effect, that all capital, once created, is permanent. In that case the introduction of a trend would still create a cycle which is basically no different from the other models (though there would be no cycle, of course, in the absence of a trend). The accumulation of new investment opportunities over time would give rise to a periodic boom, in

[1] Cf. "A Model of Cyclical Growth", paper presented to the Oxford meeting of the International Economic Association, September, 1952.

the course of which these opportunities would be exploited. The savings-investment mechanism (or shall we call it, the multiplier-accelerator mechanism) would still ensure that investment does not proceed at a steady rate, but is bunched in time. The fundamental reason, which is thus common to each variant, is that: (*a*) the output capacity of the investment goods industries is greater than the current rate of accrual of new investment opportunities—and it does not make any difference whether these new investment opportunities reflect the using-up and scrapping of old equipment or the growth of population, technical progress or anything else; and (*b*) that the economic system possesses no short-run stability at levels of investment which are below a certain maximum or above a certain minimum. Further, it is possible to introduce endless complications into the model which would tone down its sharp edges and crudities—as e.g. that investment is influenced by finance, monetary policy or rising production costs, etc.; that new inventions take the form (in part) of the creation of new industries satisfying new wants, which would explain why some investments would go on even during a depression, etc.—though it is doubtful whether in our present stage of knowledge the laborious task of working out the consequences on the model of any particular type of complication would be really worth while.

It is thus seen that all the "dynamic" models that were recently presented to the world—such as Kalecki's amended model,[1] Marrama's,[2] Hicks'[3] or Goodwin's[4] (which all contain the same basic assumptions outlined above)—are all variants of the same thing and, essentially, all consist of the superimposition of a linear trend introduced from the outside on an otherwise trendless model without altering, in any way, its basic character. Some of the authors mentioned deny this. Thus Professor Hicks thinks that "it is on the trend rate of growth that the whole cyclical mechanism depends"[5] and that in the absence of the growth in autonomous investment the cycle could not get off the

[1] *Studies in Economic Dynamics*, London, 1943.
[2] *Review of Economic Studies*, Vol. XIV, p. 34.
[3] *A Contribution to Theory of the Trade Cycle*, Oxford, 1950.
[4] "A Model of Cyclical Growth", *op. cit.*
[5] *Op. cit.*, p. 108.

floor—or, at any rate, not for an inordinately long time.[1] Similar ideas were expressed also by Mr. Goodwin.[2] This, in my view, is mistaken, and must be due to an insufficient appreciation by the authors of the properties of their own models.

Finally, there is the illuminating piece of algebra produced by Mr. Harrod[3] and Mr. Domar,[4] according to which, given the savings function, there is only one particular rate of growth that can be permanently maintained—a rate that is entirely determined by the relation of the savings function to the technical relationship between capital and output, and which is thus independent of the trend as given by the assumed rate of growth of population and of productivity. In the present context this need not detain us for long, since—as the two authors would themselves be prepared to admit—this "dynamic equation" is relevant only for determining the trend of a cycle-less economy, and not the trend rate of growth of an economy which does not actually maintain a moving equilibrium of growth, but proceeds by a series of investment booms, interrupted by slumps.[5, 6]

[1] *Op. cit.*, p. 105, note. Hicks' analysis suffers also from a confusion between the idea of "autonomous" investment (i.e. investment which is *not* responsive to changes in current output) and the idea of "new" investment opportunities due to the growth of population and technology. There is no reason whatever why these two should overlap. It is perfectly possible for there to be no autonomous investment in Hicks' sense, and yet be a trend, due to the above factors; equally there could be autonomous investment in a trendless economy. In Mr. Goodwin's recent model, on the other hand, the trend is the direct outcome of the underlying growth in population and productivity, and not of the growth of "autonomous" (in the sense of "non-induced") investment.

[2] "A Model of Cyclical Growth", *op. cit.*

[3] Harrod, "An Essay in Dynamic Theory", *Economic Journal*, March, 1939, and *Towards a Dynamic Economics*, 1949.

[4] *American Economic Review*, March, 1947.

[5] In other words—to put the matter in Mr. Harrod's own terminology—the *actual* trend resulting from successive booms and slumps reflects the "natural rate" (i.e. the growth of population and productivity) and not the "warranted rate". The warranted rate (interpreted, as it should be, as the rate of growth at *full-employment savings*) determines, *inter alia*, the duration of the boom, and also the length of the cycle, but not the trend rate of growth. This assumes, of course, that the warranted rate is higher than the natural rate—but this may be taken as a basic feature of capitalist society. If the warranted rate were *less* than the natural rate, new investment opportunities would accrue faster (in time) than they are exploited through current investment; new investment would thus become steadily more profitable; and it may be supposed that the economic system would gradually evolve the institutional conditions necessary —by raising the profit rate and the amounts set aside out of profits—for a faster rate of exploitation, i.e. for the warranted rate to rise. Indeed, the independent rôle accorded to the savings function in Mr. Harrod's and Mr. Domar's method of approach is one of the criticisms that one can make against their method as a tool for the analysis of *long-run* problems.

[6] Cf. also Joan Robinson, "The Model of an Expanding Economy", *Economic Journal*, March, 1952, pp. 52-3.

Mr. Harrod makes no claim, in his recent analysis, to put forward a complete model of the cycle; in his view, however, the "dynamic equation" provides the key to the existence of the cycle, since it is the peculiarity of any "equilibrium of steady advance" that it is surrounded, on each side, by "centrifugal forces". If it could be shown that a moving equilibrium of growth is necessarily unstable, while its static counterpart, the stationary equilibrium, is stable, this would indeed provide an inherent connection between the cycle and economic growth. But if our own analysis is correct, this is not so. Exactly, the same assumptions which make the moving equilibrium of a growing economy unstable, will produce the same "centrifugal forces" around the stationary equilibrium of a trendless economy;[1] in fact, the latter can itself be looked upon as a special case of the "dynamic equilibrium" where the equilibrium rate of growth happens to be zero. Hence there is nothing in the assumption of a positive rate of growth as such (such as the assumption of a growing population, or technical progress) that serves as an *explanation* of the tendency to oscillation that could not equally be explained without it.

The Problem of Economic Growth

So far, so good. But is this situation, from an intellectual or analytical point of view, wholly satisfactory? The trend itself is not explained; it is introduced as a datum. There can be no pretence, therefore, of these theories providing the basis for a theory of economic growth. Yet the very fact that different human societies experience such very different rates of growth—in fact, differences in rates of growth in different ages or in different parts of the world in the same age are one of the most striking facts of history—in itself provides powerful support for the view that technical invention and population growth, the two factors underlying the trend, are not like the weather or the movement of the seasons, that go on quite independently of human action, but are very much the outcome of social processes. The growth in population, in particular, is as much the consequence of economic

[1] On this point cf. also my article in *Economic Journal*, December, 1951, pp. 842-3, [above pp. 203-4].

growth as the condition of it. This is true both of situations of the Malthusian kind, where it is merely the consequence of a higher rate of survival brought about by an improvement of the means of subsistence, and where population growth is merely the passive consequence, or accompaniment, of economic growth; as also of situations where the acute shortage of labour resulting from rapid economic expansion directly stimulates the growth in numbers—as e.g. in America after the Civil War, or in England in the course of the Industrial Revolution.[1] The same is true of technical invention or innovation. Though new ideas, looked at in isolation, are the spontaneous product of the workings of the human brain, the kind of ideas that come forth, and their frequency, is very much a matter of environment. Few would deny that the greatly accelerated growth of the last 200 years, which is associated with the rise of modern capitalism, was essentially the product of forces endogenous to society; while the vast flow of technical innovation and of population growth, which accompanied this process and which alone made its realisation possible, were themselves the products of these basic social forces and not the initiating causes of it.

Where, then, are we to look for the ultimate factors responsible for the rate of economic growth in human societies? The English classical economists—if one can epitomise their views in a single word—thought they found it in Thrift. The basic factor responsible for growth and progress is the Rate of Capital Accumulation (which in the classical view was responsible for the growth of population as well as material capital); and capital accumulation, in their view, was largely a matter of the habits of thriftiness of the population—its willingness to forsake present enjoyment for future gain. At the present time we are not willing to accord such an important rôle to the propensity to save. This is not because—or not mainly because—we have learnt, since Keynes, that saving can be a positive impediment of progress, and not always a promoter of it. It is because it is now increasingly realised (perhaps also under the influence of Keynesian ideas)

[1] The fact that in Western European societies in the last fifty to seventy years the increase in population so markedly lagged behind the economic growth potentialities (owing to widespread birth control and the control of immigration) should not blind us to the force of this generalisation as a broad historical tendency.

that the proportion of income saved in the community is not so much the result of basic psychological attitudes and propensities, as of the institutional framework of society, which is itself continually (though slowly and gradually) adapted to social requirements. The re-investment of the profits of business enterprise always has been, and still is, the main source of finance of industrial capital accumulation; we are now aware that the rate of accumulation is at the same time one of the main determinants of the amount of profits that is thus available for financing.[1] Saving and Capital Accumulation therefore are in no different position from Technical Progress and Population Growth, as being one of the features that characterise progressive societies rather than as the ultimate stuff and substance which *make* societies progressive.

The most plausible answer to the question why some human societies progress so much faster than others is to be sought, in my view, not so much in fortuitous accidents, like major discoveries, or in a favourable natural environment (though no doubt all these are important as conditioning factors) but in human attitudes to risk-taking and money-making. The modern businessman or entrepreneur, as featured in economic text-books, with his distinctively speculative bent and his interests and energies concentrated on profit-making, is clearly a product of capitalist society. But the rise of modern capitalism, in turn, cannot be explained except as the result of the growth of mental attitudes which find in risk-taking and money-making the means of giving expression to the individual's *ego*. Economic speculation here trespasses on the fields of sociology and social history; and the most that an economist can say is that there is nothing in economic analysis as such which would dispute the important connection, emphasised by economic historians and sociologists, between the rise of a Protestant ethic and the rise of Capitalism.

An economy which, over longer periods, shows a relatively

[1] That distinguished pupil and critic of English classical economics, Karl Marx, has seen this clearly; and this is perhaps one of the main aspects in which, on a purely analytical level, his doctrine departs from Ricardo's. Marx, however (as far as I understand him), would regard the continuous growth of knowledge as regards the techniques of production—which goes on, in a sense, independently of society, but which shapes and determines the social relationships between men and the institutional framework of society—as the main "independent variable" of economic growth.

high rate of growth is almost certain to be one which keeps "bumping against the full-employment ceiling" fairly regularly, and with reasonable sized bumps. It is an economy which, once set upon the path of rapid expansion, does not stop expanding until it comes up against the physical limitations of capacity or labour; and it is in the strength of the incentives to overcome such limitations that we must look for the main motive force for long-term dynamic growth. Scarcity of labour, as we have argued, directly stimulates population growth—both through immigration into a particular area and through stimulating a higher effective rate of natural growth. But prolonged scarcity of labour is also the most powerful incentive to the invention and introduction of labour-saving devices,[1] just as shortages of physical capacity provide the incentive to the creation of new capacity.

The conclusion which emerges from this is that so far from the trend rate of growth determining the strength or duration of booms, it is the strength and duration of booms which shape the trend rate of growth.[2] It is the economy in which business-men

[1] The fact that the productivity of labour in American industry, over the last eighty years appears to have risen at about twice the annual rate as in European industry, could be satisfactorily explained (in my view) only through the fact that the higher expansion-mindedness of American business created prolonged periods of acute labour shortage in the economy, which stimulated both the high rate of immigration prior to 1914 and also the invention of labour-saving devices, such as the belt-line assembly, the single-purpose machine tool, the methods of mass-production generally, and far-reaching standardisation. On the other hand, the sudden slowing down of population growth in the 1920s (largely the result of the immigration laws), by reducing the trend rate of growth, must have been a major cause of the severity of the depression in the 1930s.

[2] It is essential, of course, to suppose for this that the "boom"—implying a state of affairs where production is confined by physical limitations and not by effective demand—is a matter of some duration and not only a fleeting momentary phase in the course of the cycle. I believe that to have been the case with the "strong booms" of history (though not, of course, of all booms); and one reason for my dislike of the use of the acceleration principle as a device in trade-cycle theory is that it automatically excludes this possibility. With the acceleration principle, output "bounces back as soon as it hits the full-employment ceiling" in much the same way as a tennis ball when it hits a wall. (This is because in a simple acceleration model the adjustment of the capital stock to the change in output is assumed to take up the same length of time as the adjustment of output to the change in demand. Professor Hicks has attempted to relax this condition in his own model by providing for a *series* of acceleration coefficients, thus allowing for a change in output in a given period to give rise to investment in a whole series of future periods. However, his equations make no allowance for the related fact of a back-log of desired investment accumulating as a result of the investment industries becoming clogged up with the uncompleted investment orders of the past, in which case the rate of investment ceases to have any definite relation to the rate of change in output. It is implicit in Hicks' equations that *all* the investment that is profitable to be undertaken at any given level of output has either been completed or else is in process of building as soon as that level of output is reached. Thus if we applied his equations to our numerical example on pp. 218-9.

are reckless and speculative, where expectations are highly volatile, but with an underlying bias towards optimism, where high and growing profits are projected into the future and lead to the hasty adoption of "unsound" projects involving over-expansion, which is likely to show a higher rate of progress over longer periods; while it is an economy of sound and cautious business-men, who are slow at reacting to current events, which is likely to grow at a slow rate. It is true, of course, that it is the very process of over-expansion during the boom which makes the ensuing slump inevitable. But the same cyclical force which causes booms to end after a time and be converted into slumps, also causes slumps to end and be converted into new booms; and the extent of the "over-expansion" in the previous boom influences to a major extent the degree to which the new boom surpasses the peak reached by its predecessor.[1]

This is not to suggest, of course, that the long-term trend of growth is *simply* a matter of the degree of recklessness of society's entrepreneurs. The external "conditioning factors" are still there—in the sense that there probably always is a *maximum* attainable rate of saving, a *maximum* attainable rate of population growth or a *maximum* attainable flow of new ideas. But the point is that the *actual* values of these variables, in any given society and at any given age, are not determined by their theoretical maximum values, but are capable of being slowed down or accelerated in accordance with the push or pull exerted by entrepreneurial behaviour. (Perhaps this idea could best be expressed in Mr. Harrod's terms by saying that while the warranted rate of growth and the natural rate of growth are two different things, they are not independent of each other, since the more the

above, they imply that 5 million tons of new shipping is either completed or in the process of building at the moment when the expansion reaches the full-employment ceiling. But if the capacity of the shipbuilding yards is only 1 million tons, clearly not more than 1 million tons can be in process of building at any one time; and since the maximum expansion of output is reached when shipbuilding attains the rate of 1 million tons, there must be a large back-log of desired investment by the time the "ceiling" is reached, which would prolong the boom well beyond the "two periods" allowed for in his equations.)

[1] This is particularly true of investment in investment goods industries, such as steel capacity, etc. As mentioned earlier, expectations might be responsive to changes in current profits in the consumer goods industries and yet be sluggish to respond to changes in demand in the investment goods industries. In that case not only will booms tend to be weak, but the rate of growth in output between successive booms will also be small.

warranted rate tends to exceed the natural rate, the more it will *bend* the natural rate in its own direction.)[1]

Here at last we find the inherent link between trend and cycle that we were searching for. For if the above analysis is correct, both the trade cycle and economic growth are the resultant of particular attitudes of entrepreneurs—more precisely, of the volatility of entrepreneurial expectations. As was stated at the beginning, the basic assumption underlying all recent models of the trade cycle is the dependence of long-term investment decisions on current profits: the speculative frame of mind which causes frequent and far-reaching adjustments of long-term expectations in the light of current experience. Without that no variant of the output-investment relationship, such as the acceleration principle or any other assumption, would work, and the system would tend to settle down to a stable long-term equilibrium. If expectations are responsive, but sluggish, we might get a moderate cycle, with weak booms and weak slumps and an equally weak trend. It is when expectations are highly volatile that the expansionary phase of the cycle is likely to be vigorous and sustained; that it will inevitably lead to a strong boom which will burst through the pre-existing "external frame" of the economy and carry it to a new and higher plateau. Once such a higher plateau is reached, the subsequent slump, though severe, will not mean a return to the previous depression level; it will in time produce a new expansionary process from a higher "floor", leading to a new "ceiling".[2]

[1] Mr. Harrod seems to me to have drawn the wrong moral from his own analysis in suggesting that an excess of the warranted rate of growth over the natural rate will make "the economy to be prevailingly depressed", although "at first blush, one would suppose it to be a good thing that the line of entrepreneurial contentment should be one implying an attempt to push forward always at a greater rate than fundamental conditions allow" (*Towards a Dynamic Economics*, p. 38). Our own analysis would suggest that Mr. Harrod's "first blush" was right, and his subsequent analysis was wrong. What he did not allow for was that the "fundamental conditions" determining the natural rate of growth are not determined by Heaven—they are pliable (within wide limits) and can be pushed outwards or pulled inwards by the endogenous forces of the economic system.

[2] At first sight the proposition that higher volatility of entrepreneurial expectation as such promotes a higher trend rate of progress might appear paradoxical, since the stimulus provided by the greater optimism of the boom would appear to be offset by the greater pessimism of the slump. But my contention is that the effects of the two are not symmetrical. The worst that slump pessimism can do is to interrupt the investment process altogether. But if the tide of boom optimism carried the economy to a higher plateau of productivity and of standards of living, the cessation of investment in the subsequent slump will not mean a return to the floor of the previous

The same forces therefore which produce violent booms and slumps will also tend to produce a high trend-rate of progress; though the connection between the two is far too complex to be reducible (at present) to a simple mechanical model. And Schumpeter's hero, the "innovating entrepreneur", whom we dismissed so summarily and rather contemptuously at the beginning, is found, after all, to have an honourable place, or even a key rôle, in the drama—even though we prefer to endow him with a rather more variegated character. He is a promoter, a speculator, a gambler, the purveyor of economic expansion generally, and not just of the "new" techniques of production.[1]

slump—the floor will have risen—while the rise in the full-employment ceiling as between successive booms is itself largely conditioned by the tide of optimism in booms. Of course, if entrepreneurial attitudes are in themselves asymmetrical—if there is a consistent bias towards optimism—the stimulus to progress will be even stronger than if they are merely volatile.

The above view receives strong support from (and in fact is greatly dependent on) Mr. Duesenberry's analysis of consumption behaviour and saving in *Income, Saving and the Theory of Consumer Behaviour*, according to which once a higher standard of consumption is reached in the boom, the consumption function shifts, and the subsequent fall of consumption in the slump is less than it would have been otherwise. Equally, it supports the view that the maintenance of consumption in slump periods by means of anti-cyclical fiscal devices, unemployment pay, etc., must tend to accelerate the trend rate of growth of the economy.

[1] The conclusion that in a capitalist society—i.e. in a society where investment decisions are made by a multitude of entrepreneurs in the light of profit-expectations—the trend rate of growth is likely to be all the greater the more powerful are the cyclical forces, is not to be taken to imply that progress must necessarily take the form of fits and starts, whatever the institutional arrangements of society. If investment were centrally planned and the consumption function continually adjusted to secure full employment (given the planned rate of investment) there is no reason, in theory, why progress could not take place at an even rate.