

EXPLAINING ECONOMIC GROWTH

By M. FG. SCOTT

Read 11 November 1986

Introduction

I AM going to contrast two different ways of setting about the task of explaining economic growth. Both use a simplified model of the economy. The first is the neoclassical growth model based on an aggregate production function in which capital and labour are the inputs. Capital consists of capital goods, and investment then involves increasing the quantity of such goods. Theoreticians have referred to capital as if it consisted of corn, steel, machinery, or tractors, or even Meccano sets or jelly. I shall call this *orthodox* growth theory, and will defend the adjective shortly.

The second way is one which I have developed over the last ten years or so in which the capital stock is not of central interest, although investment is. In a completely static economy, there is no investment, only maintenance expenditure, and nothing changes. In a growing economy, economic arrangements change, and investment is defined as the cost, in terms of consumption foregone, of changing economic arrangements. In default of a better title, I shall call this simply my growth theory, although I am quite ready to be told that someone else expounded it long ago.

Time does not permit me to set out either theory fully or fairly. You will, in any case, be familiar with orthodox theory. All I can hope to do is, first, to convince you that that theory is profoundly unsatisfactory. While that is hardly news, some of my reasons for dissatisfaction may be new. Secondly, I hope to start you thinking along similar lines to my own. Following through the definition of investment which I have just given has led me to some unexpected conclusions. I will not be able to reach them all in this lecture, but perhaps I can persuade you to begin the journey towards them.

In 1958 Kaldor¹ suggested six 'stylized facts' as a starting point

¹ See his 'Capital Accumulation and Economic Growth' in Lutz and Hague (1961), pp. 178-9.

for the construction of theoretical models of growth. In 1969, Solow¹ endorsed these. Although they have since been criticized, notably by Hacche,² my own investigations suggest that a modified version of them does provide a valuable way of summarizing aggregate economic behaviour over quite long (but *not* indefinitely long) periods. In the modified version, the following four magnitudes for the *non-residential business sector*³ of the economy are all constant over selected periods of ten to thirty years, and the economy is then said to be in a state of *steady growth*:

g, the growth rate of output;

g_L , the growth rate of quality-adjusted⁴ employment;

s, the ratio of gross investment to the gross domestic product of the sector, both measured at current prices;

and λ , the ratio of income from employment⁵ to the gross domestic product of the sector, both measured at current prices.

By 'constant' I mean, of course, that the *trends*, when cyclical fluctuations are removed, are constant, these trends being approximately zero for s and λ .

A sketch of orthodox growth theory

The most important characteristic of orthodox theory which distinguishes it from other theories of economic growth and, in particular, from my theory, is the assumption that technical progress is independent of investment.

The process of economic growth is pictured as follows. At any given time entrepreneurs are aware of a given technology. This has been described as a 'book of blue prints'⁶ and consists of knowledge of a whole set of techniques for producing output. In the light of their economic circumstances (prices, existing stocks of land, labour and capital, market situation, etc.) entrepreneurs

¹ Solow (1970), p. 7.

² Hacche (1979), ch. 15.

³ i.e. the GDP *less* the output of public administration and defence, health, education, and the services of dwellings.

⁴ An index of quality-adjusted employment is an attempt to weight numbers employed by their relative marginal products. Allowance is generally made for hours of work, age- and sex-composition, education, and the transfer of labour from agriculture and unincorporated businesses generally to higher paid sectors of the economy.

⁵ Including that part of the income of the self-employed deemed to be income from employment, and before deduction of direct taxation and national insurance contributions.

⁶ Robinson (1962), p. 116.

select the techniques they will use to produce outputs. As time passes, three main changes occur.

First, the labour force may change. In the simplest theories, labour is homogeneous, so changes in the labour force consist simply in an increase or decrease in numbers. It is perfectly possible, however, to allow for changes in the quality of the labour force, at least approximately, and still to retain the highly convenient simplification of a single magnitude, L , to represent it. This can be done by the well-known device of weighting the quantities of different types of labour in proportion to their marginal products before adding them up. Each unit, appropriately weighted, has then the same (marginal) productive power, and so the total of all weighted units is rendered homogeneous in at least that respect. By this means the theory can include the effects on production of such things as changes in the age- and sex-composition of the labour force, or changes in the amount of education or training it has received. Denison is the leading exponent of this method.

The second main change to occur is that the quantities of different capital goods available may change as a result of two processes. On the one hand, new capital goods are produced. On the other, old capital goods are consumed, or partly consumed. 'Capital goods' include stocks of finished and semi-finished goods and raw materials, as well as things like machines, vehicles, buildings, and construction work generally. They do *not*, however, include 'blueprints' or 'knowledge'. Just as it is highly convenient to combine different types of labour together so as to form a single number, L , standing for the whole labour force, so it is convenient to form a single number, K , to stand for all capital goods taken together. It seems that the best way to do this, in principle, would be to weight the quantity of each type of capital good in proportion to its marginal product, just as for labour. Surprisingly, and despite statements to the contrary, no one (so far as I know) has yet done this, even approximately. I return to this later, as the failure is an important criticism of empirical applications of orthodox growth theory, as well as being a serious mistake in many expositions of the theory itself. As it is not, however, my immediate concern, I shall simply assume for the present that the weighting has been correctly done.

The third main change to occur is in technology. More techniques become available from which entrepreneurs select those which, in their circumstances, yield higher output for given inputs, or else lower inputs for given outputs, or some combination of these. Orthodox theory does not itself seek to explain why or

how these new techniques become available: the process is exogenous to the economic system so far as the theory is concerned. There is a considerable literature on the subject, but the assumption of exogeneity has remained intact—not surprisingly, since changing it would profoundly alter the theory and its main conclusions, as we shall presently see. The implication is that technical progress cannot be appreciably speeded up, or slowed down, as the result of economic activity and, notably, as a result of increasing or reducing any kind of investment expenditure.

The simplest, and so most widely used, orthodox theory employs the following aggregate production function:

$$Y = f(K, L, t) \quad (1)$$

Here, Y is total real output or income, K is the total real capital stock, L is the total labour force employed, and t is time. Y , K , and L must all be interpreted as some sort of index numbers which combine heterogeneous quantities by suitable weights in the manner just described. Increases in K , L , or t each increase Y . The increase which results purely from the passage of time, K and L being constant, is due to technical progress.

An apparent advantage of (1) is that it explains both the *level* of output at any given time and *changes* in that level from one time to another. I question this quite staggering claim later.

Is neoclassical growth theory orthodox?

Before proceeding with the argument, I must justify my assertion that the neoclassical growth theory I have just described is the *orthodox* growth theory.

I offer two criteria. First, there is the ability to survive in textbooks for students. Their writers are forced to select from the welter of theories available, and what they select is what most students learn. Examination of a fair number of textbooks does show that virtually all make use, in their treatment of growth, of the concept of the production function with output depending on inputs of capital and labour, with diminishing returns to each of these factors, and with technical progress shifting the function through time. This is the centre-piece of neoclassical theory, and no other theory is so well represented in the textbooks.

As well as this textbook test we can apply the test of empirical application. If a theory is successful it will be used to explain historical growth and, possibly, to predict future growth. A theory which cannot be, or is not, used for either of these purposes is

ultimately valueless, if we regard economics as having any pretensions to being a science. The main empirical application of orthodox theory has been in growth accounting, which has not only been frequently done,¹ but has also led to some important conclusions which I refer to later. By comparison with this, I know of no substantial body of empirical work on growth which is based on any other modern theory.

Many writers have questioned the assumptions underlying orthodox growth theory. Thus three of the economists who perhaps did most to launch the theory initially, Meade, Solow, and Swann, all expressed their doubts about it. Nevertheless, no rival theory has yet succeeded in taking its place, and orthodoxy marches triumphantly on. Let us now see how it explains the stylized facts.

The problem which orthodox theory has to solve can be put as follows (I shall be very brief, as I am treading ground that will be familiar to all of you). According to Kaldor's stylized facts, in steady growth the capital/output ratio is constant, which means that capital, K , grows as fast as output. However, in a progressive economy labour productivity grows, which means that quality-adjusted labour, L , grows more slowly than output. It follows that the ratio of capital to labour, K/L , must be rising. One would then expect that, with diminishing returns, the rate of return to investment would be falling, but another of Kaldor's stylized facts is that the rate of return is constant. How can that be explained? The explanation offered is that technical progress is of a particular, labour-augmenting, kind called Harrod-neutral technical progress. I do not myself find this explanation persuasive, but as that is not the main criticism I wish to make of orthodox theory, I will spend no more time on it. I will, however, mention another feature of orthodox theory which others besides myself have thought surprising, and that is what might be termed the iron law of growth. This is that, in the long run, and no matter whether one invests one, ten, or fifty per cent of total output, the rate of growth is the same, and is equal to g_L plus the rate of Harrod-neutral technical progress.

¹ The leading exponent is E. F. Denison (1962), (1967), (1974), (1976), (1979), and (1985). Other important works are: the earlier studies of Fabricant (1954), Abramovitz (1956), and Solow (1957), all emphasizing the great importance of technical progress; the historical studies of France by Carré, Dubois and Malinvaud (1976), of Japan by Ohkawa and Rosovsky (1973), and of the United Kingdom by Matthews, Feinstein and Odling-Smee (1982); and see also the survey by Kennedy and Thirlwall (1972).

It is clear that orthodox theory attaches great importance to the rate of technical progress, and empirical support for that importance is provided by the growth accounting studies to which I have already referred. Some of the earlier studies¹ exaggerated the contribution of technical progress by not allowing for quality improvement of the labour force which should be included in g_L , and will be included if we weight different members of the labour force by their relative marginal products. Denison has made careful allowance for such quality improvements, and has also assumed that a substantial part of growth is due to economies of scale, so that the respective contributions of both capital and labour are larger on that account. He has also allowed for the effects of a number of other factors on growth, such as changes in the legal and human environment, changes in the weather, labour disputes and changes in the intensity of demand—all of which, however, are rather minor over long periods. In his latest study of US growth, the contribution of his residual, ‘advances of knowledge and n.e.c.’, to growth was about twice as big as that of capital over the whole period surveyed, 1929–82.² Although Jorgenson and Griliches³ made estimates using the same basic methodology which appeared to show that there was virtually no residual at all for the US private domestic economy over the years 1945–65, these were subsequently (and in my opinion successfully) challenged by Denison, and Jorgenson and Griliches revised their estimates so as to show a much more substantial residual.⁴

Having briefly reminded you of the main tenets of orthodox

¹ e.g. Solow (1957).

² See Denison (1985), Table 7-1, p. 107. The figures all refer to contributions to the growth rate of national income in the non-residential business sector of the USA. For the period 1929–82, the contribution of capital is put at 0.38 percentage points per annum, and that of advances in knowledge and n.e.c. at 0.86 percentage points per annum. The contribution of economies of scale is additional to both of these and is put at 0.34 percentage points per annum. If it were allocated to capital and labour in proportion to their contributions as calculated assuming constant returns to scale, the contribution of capital would rise to 0.45 percentage points per annum. It must be said, however, that Denison’s allowance for economies of scale is, at best, a well-informed guess. The relative importance of the residual is appreciably greater in the sub-period 1948–73, and far greater still in 1929–48. The residual is negative in 1973–82.

³ Jorgenson and Griliches (1967).

⁴ Jorgenson and Griliches’s original (1967) article with Denison’s ‘Examination’ of it were reprinted in US Department of Commerce (1972) along with Jorgenson and Griliches’s ‘Reply’, Denison’s ‘Final Comments’, and Jorgenson and Griliches’s ‘Final Reply’.

growth theory, and of the way in which it explains the stylized facts of growth, I now turn to my main criticisms of it.

The idea of reduplication

The simplest expositions of orthodox theories explicitly assume that 'capital' consists of homogeneous capital goods (e.g. 'corn', 'steel', 'tractors'). Implicit in this assumption is the view that actual investment in the real world can be regarded as the *reduplication* of existing assets. One is adding more to existing stocks of machines, vehicles, houses, factories, etc., and what one is adding is essentially the *same* as what is already there. For production function theories this idea is quite fundamental. It is intimately related to the idea of exogenous technical progress, since any improvements in capital goods are attributed, not to investment as such, but to this mysterious outside force.

The idea of reduplication is needed to justify the following three important characteristics of production functions.

First, most expositions of production function theory assume *constant returns to scale*, although some allow for increasing return to scale. The assumption of constant returns to scale is only plausible if applied to equal proportionate changes in homogeneous inputs. Thus it is plausible to suppose that if twice as many men work with twice as many spades on twice the area of land they will produce twice as much. But if we give twice as many men on twice the area of land some tractors instead of spades, no simple prediction about the change in output seems possible, even if we are told that the tractors cost twice what the spades cost.

Secondly, all production function theories assume that there are *diminishing returns to the capital stock*. Again, the implicit idea is that additions to the capital stock mean more of the same (more spades, not tractors instead of spades). If that were so, then indeed one could expect that the marginal product of extra spades would diminish as more were added to a given number of men working a given area. But if, when one invests more, one provides, first, better spades, and then tractors, and then better tractors . . . etc. then it is far from clear that more means worse. Does the existence of a past history of massive investment (i.e. a large capital stock) imply that investment *now* has a lower return than if there were a past history of meagre investment (i.e. a small capital stock)? Is the rate of return in Japan, say, lower than in the Sahara? Or was the rate of return in Japan in 1980 lower than in 1880? Or would the rate of return to *further* investment in the Sahara be lower if, before that further investment were undertaken,

a massive investment programme had been undertaken which succeeded in developing its water resources, covering it again with vegetation, changing its climate, and making the desert bloom? Put that way, diminishing returns to the capital stock cease to be so obvious.

Thirdly, production function theories assume that technical progress shifts the function through time. This is bound up with the idea of reduplication because, in default of technical progress, reduplication is all that would occur. Hence the distinction between movements along the function and shifts of the function depends on whether reduplication is or is not occurring. In my theory this distinction disappears. Since investment changes economic arrangements it seldom is reduplication, and nothing crucial depends on whether or how far it is.

It should be clear from these three examples that the idea of investment underlying production function theories is quite different from the idea of investment as expenditure incurred in changing economic arrangements. It is only a very special kind of change, namely, reduplication of existing capital goods, which is regarded as investment by production function theories. Since most investment is not of this kind, can one rescue the idea of reduplication in any way? The only way that I know of is the following artificial device which has been best explained by Denison.¹ Suppose the capital stock, defined in a way to be discussed, is £K. Suppose net investment in some particular period is £ΔK. Then one *assumes* that the effect of this investment on output is *as if* each and every item in the capital stock had been increased in the proportion ΔK/K. If the actual increase in output (abstracting from changes in the labour force) is greater than this (and the presumption is that it will be greater), then the difference is not attributed to investment but to technical progress, including the catching-up of old-fashioned techniques with best-practice

¹ See, for example, the explanation of his basic method in Denison (1967), pp. 33-4. After first noting that he assumes constant returns to scale in this part of his calculations, and allows for increasing return separately, at a later stage, he asks, 'What fraction of the increase in real national income that would result from a 1 per cent increase in all factors of production is obtained from a 1 per cent increase in only one factor or group of factors?' He answers, 'Suppose, for example, that the factor is labour and that labour earns 80 per cent of the national income . . . A 1 per cent increase in the quantity of every type of labour in use will then be equivalent to an increase of 0.80 per cent in all types of input'. While Denison's illustration refers to labour, he applies the result to capital as well, and so he must be treating investment as if it were bringing about a uniform percentage increase in every type of capital.

techniques (i.e. changes in the lag of application of techniques, as Denison calls it).

In this way we can preserve constant returns to scale and diminishing returns to the capital stock. However, we have preserved them tautologically: there is no way of testing, empirically, whether they exist. They are simply assumptions, no more and no less. The question then to be answered is whether this is an illuminating and interesting way of describing reality, and whether the residual effect of 'technical progress' corresponds to anything interesting. I rather doubt it. There is no reason to suppose, for example, that technical progress, so defined, measures the effect of research and development expenditures. Indeed, I cannot think what it measures except (tautologically) the difference between an actual increase in output and a purely hypothetical increase which is based on a set of definitions which I can see no reason for using.

No investment, no growth

Some may object that the preceding argument misinterprets production function theory. The assumptions of a homogeneous capital stock, or of investment as a uniform proportionate increase in every item in a heterogeneous capital stock, are *not* the only assumptions which can be or are made. On the contrary, it is possible to cover the case of a heterogeneous capital stock in which the different items are increased in different proportions by the usual device of weighting each increase by its marginal product. Jorgenson and Griliches are one of the clearest expositors of this method. Since I use it so far as labour inputs are concerned, one may reasonably ask why it cannot be used for capital inputs as well? The answer is that it can, but that in practice, and so far as I am aware, it never has been, and certainly not by Jorgenson and Griliches. The result of using it is, inevitably, to eliminate the contribution of 'technical progress' to growth altogether.¹

Let us first consider how in practice changes in capital inputs are measured by those who have attempted to explain or account for economic growth. Two measures of the capital stock have been used, namely, the gross capital stock at replacement cost new and the net capital stock. I take these in turn.

¹ This is what Jorgenson and Griliches (1967) indeed claimed to have done (or nearly done—a small residual remained). However, as already noted, their demonstration was shown to be faulty by Denison (see p. 274, n. 4), and their argument was not that given below. The objections which follow to the capital stock measures used by practitioners in this field apply to Jorgenson and Griliches's capital stock as well. Hence they reached the right conclusion for the wrong reasons.

The gross capital stock at replacement cost new is generally calculated by the perpetual inventory method. An estimate is made of the average life, θ , of each type of fixed capital asset, such as machinery. Gross fixed investment is estimated for θ years back from the year, t , for which an estimate of the stock is required. Investment before that is deemed irrelevant, since none of the assets then created will have survived to year t . After converting these investment figures to year t prices (or to the prices of whatever base year is chosen for constant price estimates) by means of an appropriate index of machinery prices, the gross investments over θ years are simply added together to give that component of the stock. This is added to other components, similarly calculated, for other types of asset with different lives. Finally, stocks of goods and work in progress are added to give the total capital stock (for some mysterious reason, several authors omit this final step, as if such stocks had no marginal product). Refinements are possible, in which assets of a given type are assumed to have lives which are not exactly the same, but are distributed around the mean—but we need not concern ourselves with this here.

This procedure could be justified in terms of the theory only if the marginal product of each asset was, throughout its life, directly proportionate to its original cost (brought to a common price level by index numbers of asset prices). This is unlikely to be the case. At the end of θ years, the asset is going to be scrapped, not, in the majority of cases, because it has physically worn out, but rather because the quasi-rent it earns has fallen to zero, or close to zero. In other words, its economic life is determined by economic factors (namely, competition from newer and better assets), and not by physical ones: by obsolescence rather than by decay.¹ That being so, the clear implication is that the asset's quasi-rent must fall over its lifetime, eventually becoming approximately zero. This is so despite the fact that the physical output, and other physical inputs, associated with the asset may all remain approximately constant over the whole of its life. Because relative prices change (typically, real wages rise), the asset's quasi-rent will typically decline, and so will its marginal product.

Some may feel that, if the *physical* inputs and outputs remain unchanged, then the *relevant* marginal product, so far as the

¹ See Barna in Lutz and Hague (1961), p. 85, 'it is obsolescence rather than wear-and-tear which is the dominant cause of mortality-homicide to make room for a new favourite, rather than natural death'. See also Kuznets in Conference on Research in Income and Wealth (1951), p. 65, and Kuznets (1961), p. 396.

production function is concerned, has remained unchanged. This seems indeed to be Denison's belief. Yet it seems to me self-evident that if, to take the extreme case and the simplest assumptions (perfect competition), an asset is scrapped whose marginal product, equal to its quasi-rent, is zero, the effect on output must be zero. The other factors of production associated with the asset must then be able to go elsewhere and produce as much as they were earning when they worked with the asset, and this must equal the whole of the output they then together produced. The economic contribution of the asset is then zero, and scrapping it does not reduce total output. Yet it *does* reduce the gross capital stock, as conventionally measured. Hence it is patently inconsistent with the production function. The gross capital stock at replacement cost new cannot be the correct way to measure K if we want to use that function.

Let us now consider the other measure of K which is commonly used, namely, the net capital stock. This can also be estimated by the perpetual inventory method, but now each addition to the stock, after correcting for price changes to the base year, is written down over its lifetime by some depreciation formula (e.g. straight line, declining balance, double declining balance, etc.). The exact formula is not important for the argument which follows. The general effect is to reduce the value of any asset by an amount which increases with the age of the asset. An asset which is about to be scrapped should, in principle, have been reduced to zero value. At first blush, therefore, it seems that this method should meet most of the objections made to the gross capital stock, since the depreciated values of different assets should be *roughly* proportionate to their marginal products, although the correspondence *would* only be rough, and would depend on the formula used and on the way in which the asset's quasi-rent changed over its life. Unfortunately, there remains a serious objection to this measure, which is quite separate from the question of exact proportionality of depreciated value to marginal product for *existing* members of the stock. Instead, it relates to the method used to calculate net additions to the stock. With the net capital stock, the net addition to the stock equals gross investment minus depreciation.

The point I now wish to make really requires more explanation than I can give here. I have given a fuller one elsewhere.¹ The point is that depreciation, properly defined and distinguished from maintenance, is not a net social cost. Maintenance should be defined, and indeed is to a reasonably close approximation, as the

¹ See my 'Maintaining Capital Intact' in Collard *et al.* (1984).

physical maintenance of capital assets. Depreciation is then their decline in value due, and due only, to relative price changes. This is the main content of obsolescence. In a progressive economy it is essentially the result of rising real wages. Income is then transferred from capitalists to workers, and workers benefit from *appreciation* by exactly the same amount as capitalists lose by *depreciation*, so that there is no net loss to the economy as a whole. It should then be clear that net investment for society as a whole is (approximately) equal to *gross* investment as conventionally measured, and not to gross investment minus depreciation. Hence, if the marginal products of existing members of the capital stock are, let us say, on average r times their net capital values, the marginal product of the *additions* to the stock are much more than r times *net* investment. Instead, they are more like r times *gross* investment.

An analogy with the estimation of labour's marginal product may be helpful. Suppose we have a fixed quantity of land on which is employed a quantity of labour, L , at a wage-rate w , which we assume equals its marginal product. Now let there be a small addition to the labour force ΔL which causes the wage-rate and marginal product to sink by Δw . The contribution of the extra labour to output is measured to a close approximation by the new (or old) wage-rate multiplied by the increase in employment, i.e. by $w \cdot \Delta L$. One would not think of deducting from this the fall in the wages earned by the existing labour force, $\Delta w \cdot L$, which is merely a transfer of income from wage-earners to landowners. There is no net *social* cost involved in that, since the landowners benefit as much as the wage-earners lose. My argument is that depreciation on capital assets is loosely analogous to the wage-earner's loss and is offset by appreciation accruing to workers (like the extra rent accruing to landowners), leaving the net contribution of extra capital equal to something like r multiplied by gross investment.

The upshot of the argument so far is that neither of the two measures of the capital stock which have been used in practice correctly weights capital inputs by their marginal products. Each biases the estimate of the contribution of capital to growth downwards by a very large amount. In each case, the contribution to growth is proportionate to (say r times) the absolute growth in the stock. With the gross capital stock, one measures the absolute growth in the stock by gross investment minus scrapping, and scrapping is equal to the original cost of the scrapped assets brought up to some base year price level. As we have seen, since

scrapped assets contribute nothing to output,¹ this deduction is incorrect. With the net capital stock, one measures the absolute growth in the stock by gross investment minus depreciation. Again, this deduction is incorrect since depreciation is essentially a transfer of incomes due to relative price changes. The conclusion is the same in both cases: the contribution is proportionate to gross investment minus nothing.²

If the contribution of capital to growth is proportionate to gross investment, is anything left for technical progress? At first blush, this would seem to be a matter for empirical calculation, but I want to make it a matter of definition. By defining investment as the cost of changing economic arrangements I attribute to investment and labour force growth all the changes in output which occur, with two major exceptions. These are exogenous changes such as the weather, earthquakes, etc., and changes in capacity utilization, typically due to the ups and downs of the trade cycle. Putting those aside, let us consider first the simplest case in which there is no change in labour input. Unless there is a change in economic arrangements, we shall then have a static economy with no change in output at all. Technical progress cannot take place without a change in economic arrangements. The next step is to assert that changes in economic arrangements always cost something. Even if a few examples can be given which contradict this assertion, they are of such minor importance that they can be neglected. The conclusion is, then, that the whole change in output must be due to the change in economic arrangements, and hence to the investment which is the cost of making them. Of course, it may not be a simple matter to decide exactly what this cost is. One change leads to another, or, indeed, to several others, and so it is not easy, and may even not be possible, to trace the consequences of an investment forwards in time, nor to trace the costs incurred to make some change backwards in time. All this means is that tracing cause and effect is very difficult in practice. I therefore represent complex reality by a much simpler model of it. In that model, the costs incurred to make each change are well defined, and it is incurring those costs (i.e. investing) which is deemed to have caused those changes.

Changes in labour input must now be taken into account, and

¹ While I believe this to be generally true in normally prosperous times, in recent years of severe recession assets still capable of contributing to output may have been scrapped in large quantities.

² This argument is essentially the same as in Scott (1976) and, more fully, Scott (1981).

here I follow essentially the same procedure as everyone else. I assume that employers ensure that changes in employment are pushed to the point at which their contribution to the net present values of the firms employing them is zero. I can then assume that the contribution of labour input growth to growth of output is measured by a simple function of the wage.¹ Subtracting this contribution from the actual increase in output leaves the contribution of gross investment.

The above procedure leaves nothing for a separate contribution by 'technical progress'. This does not mean that technical changes are not occurring. On the contrary, they are continuously occurring as every investment changes economic arrangements. What it *does* mean is that the attempt to separate the contribution of technical change is abandoned. Indeed, with my definition of investment separation is strictly meaningless. Every investment changes something, and so is virtually bound to necessitate totally new knowledge in *some* degree. What is unclear is just how one can separate the additions to output due to totally new knowledge from those due to the application of already known techniques and both, in turn, from those which would have occurred had the investment reduplicated existing assets. Furthermore, what is the point of trying to make this separation? If investment is inevitably a package, one of whose contents is an increase in knowledge, why try to calculate its contribution to growth as if this content was missing? Other writers have cast doubt on the attempt at separation,² but some have cheerfully continued with models in which the separation is essential, and which have led them to the paradoxical conclusions noted earlier.

To guard against misinterpretation, it must be emphasized that I am not *merely* asserting that all technical change has to be embodied. Standard vintage theories assume embodiment, and yet lead to much the same conclusion as the orthodox theory I have criticized. The reason they do so is that the rate of improvement from one vintage to the next is exogenously given, and is independent of investment. That is not the assumption I wish to make, because it carries the following absurd implication. Imagine that 100 years ago in a closed economy (or the whole world, say) all investment ceased, and also all population growth.

¹ In a static economy with perfect competition wages equal marginal products, but in a growing economy even maintaining the assumption of perfect competition this is no longer true for reasons which require more space than is available here.

² See, in particular, Kaldor (1957), pp. 595-6.

For the next 100 years, capital assets were all maintained, so that each and every output remained constant. Arriving at the present after a century of stagnation, we start to invest again and what do we find? According to standard vintage theory, the new machines available will be capable of producing jet aeroplanes, lasers, micro-computers, the whole array of modern drugs, and all the rest. Silently, without the need for any intervening investment, technical progress will have gone on, and the modern vintages actually available will miraculously be available in the hypothetical present too.

This is all wildly implausible. It seems more plausible to liken economic progress to a journey with investment as the cost of travel. If the traveller pauses on the journey, he no longer incurs the cost, but neither does he progress and, when he starts off again, it is from where he stopped, not miles further on. One might try to preserve the standard vintage theories as an approximation to reality by maintaining that *some* investment is required to develop new vintages, but that it is very small. The idea that R & D expenditures, plus costless inspirations, are the source of all productivity growth is, indeed, widely held and must underly production function theories as well as vintage theories. It is, however, not very plausible. It would imply that the rate of return to such expenditures was very high, which makes one wonder why they are not increased.¹ It would also neglect the extent to which people learn from investments other than R & D expenditures. Acquisitions of knowledge and learning from experience are undoubtedly of great importance in explaining economic growth. However, I doubt whether knowledge and experience can be acquired other than by changing economic arrangements and seeing what results, and, if one cuts back the pace of change by cutting back investment, one must cut back the rate of acquisition of useful knowledge as well.

Suppose we grant that the contribution of capital to growth is proportionate to gross investment, without any deduction, is it possible to rescue production function theory? Let us assume that the rental of every capital asset equals r times its value in the net capital stock, K , so that gross profits are $P = rK$. In any short

¹ See Scott (1981) p. 214, where I estimated that if Denison's residual growth rate of 1.41 per cent per annum in the US non-residential business sector from 1948 to 1973 which is attributed to 'advances of knowledge and n.e.c.' (Denison (1979), p. 92) had all been due to R & D expenditure during that period its average rate of return would have had to be about 50 per cent per annum (and about 2,300 per cent per annum if only basic R & D expenditure is included).

interval, the capital stock increases because of gross investment, S , which is a *quantity* change and so is relevant to, and a cause of, increases in output. It also falls because of depreciation, D , which is due to relative price changes, and so has no direct relevance as a cause of changes in output. Rather, it is a *consequence* of investment and rising real wage rates. The (quality-adjusted) labour force also increases by ΔL from its level, L . Then, if Y is real output and w is the average wage, we might put:

$$\Delta Y = r \cdot S + w \cdot \Delta L \quad (2)$$

This would give us a growth accounting equation which is in several respects similar to the orthodox one. There are, however, some very important differences.

First, there is no separate term for technical progress. If we can integrate the above equation to form a production function, that function is not shifting through time because of exogenous technical progress.

Secondly, integrating the term for capital does not give the net capital stock, but cumulative gross investment. To get the net capital stock we would have to integrate $(S-D)$, net investment, but that is not the relevant magnitude for explaining growth. It is then not possible to write the level of output as a function of K and L as in (1), with K as the net capital stock.

Thirdly, however, if we substitute cumulative gross investment for K in (1) we encounter the following problem. How far must we go back? The perpetual inventory method places a limit on the period for which gross investment statistics are needed, which is the life of the longest-lived assets in the inventory.¹ If some other basis of asset valuation were used, such as insurance values or book values, we would not need to go back in time at all. Cumulative gross investment, however, goes back in principle without any obvious limit. Admittedly, because of exponential growth, the amount of investment in earlier years becomes small, and could be neglected for that reason. However, this question suggests a more fundamental one which confronts every historian. Can one explain the present without going back an indefinite distance into the past? The production function appears to make the staggering claim that one can, but the strict necessity of substituting cumula-

¹ There are, however, some grounds for unease. 'Land' is usually left out on the argument that it does not change appreciably in total. The age of some items in the stock (roads, hedges, cleared fields, drains, and even some buildings) can be much greater than the maxima assumed, and what 'age' means for some is unclear.

tive gross investment for the net capital stock suggests that that claim is false. Would it not be better to say that we can explain how the present has been reached from some point in the past, and that that point has to be chosen on grounds of practicality, such as the availability of data and of time and knowledge to provide the explanation? Should we not admit that we cannot really explain levels, but, at best, only changes from one level to another?

Fourthly, if we make output a function of cumulative gross investment and employment that function will not have the usual properties. In particular, and as already noted, there is no reason to suppose that the marginal return to investment will be smaller the larger is cumulative gross investment for given L . This in turn implies that steady growth in the long run may well depend on the savings ratio, so that the iron law of growth has rusted away.

Finally, we may note that if both gross investment and the quality-adjusted labour force are regarded as inputs, and if there is no separate contribution from technical progress, 'total factor productivity' cannot be measured in the conventional way. However, economic efficiency can still be measured by the rate of return to investment, or by suitable application of an equation similar to (2) to changes in outputs and inputs.

A sketch of my theory of growth

Enough of orthodoxy. I will devote my remaining time to sketching out an alternative theory of growth in which investment is viewed as the cost of changing economic arrangements, all changes are lumped together with no attempt being made to distinguish movements along production functions and shifts of those functions, there is no separate technical progress, and the attempt to explain the level of output in terms of the existing state of the world is abandoned. Instead, the level is taken as a datum and the theory explains how it changes. The level at time, t , can still be explained in terms of the level at some earlier time, t_0 , and the changes made in between, but that leaves the level at t_0 as the datum.

At a given time, a given firm will have a particular organization, productive assets, and labour force. It will be buying a stream of current inputs consisting of materials (including semi-manufactures and services) of various kinds, which it then transforms using its assets and labour force into a stream of current outputs, which are sold. The firm could, in principle, keep its organization unchanged, maintaining all its assets and also its labour force, so that retirements are balanced by recruitments. It

could then transform a constant stream of inputs into a constant stream of outputs. Its value-added, or output, at constant prices would then also be constant.

However, prices may, and in general will, be changing. Real¹ wage-rates will generally be rising in a progressive economy. Then, even if the real prices of the firm's material inputs and outputs do not change (so that it is, in that respect, representative of the whole economy), and the real value of its output then remains constant, a larger and larger fraction of that output will accrue to labour, and real gross profits will fall. This fall in real profits will cause the firm's assets to depreciate in real value.

To offset this depreciation, the firm needs to invest, and thereby to change its organization. Investment will (if wisely done) increase real gross profits, and depreciation could indeed be defined as the amount of investment required just to offset the forces making for a fall in profits, so that a firm investing exactly its depreciation will experience constant real gross profits. Typically, the firm will be investing at a faster rate than this, so that real gross profits will no longer be constant but growing. To maximize the value of the firm to its owners, the management must push investment to the point where its marginal real rate of return equals their real rate of discount.

This shows the way in which the decision about how much to invest can be taken. There is only one other decision which the firm in steady growth in my model is called upon to make, and that relates to employment. In choosing its investments from the opportunities confronting it, the firm can select those which increase employment more or less. For a given investment expenditure, a bigger increase in employment must be accompanied by a bigger increase in output in order that the higher wage bill is covered. The *extent* to which increased output is required for a given increase in employment depends on the real wage. Investment can be employment-and-output-increasing, but it can equally be employment-reducing, with output growing less, or perhaps even falling where a lot of labour is saved. In short, the firm has to decide on the labour-intensity of its investment projects, and that decision has implications for the rate at which both employment and output increase, and depends on the current level of wages.

A firm which is in steady growth is then continually changing its

¹ The term 'real' here means that wage-rates etc. are measured in terms of a *numéraire* which is the final output of the economy, which in my view means consumption. See Scott, 'What Price the National Income?' in Boskin (ed.), (1979).

organization, at a cost which is its investment expenditure. It is continually confronted by investment opportunities from which it selects only some. It makes two decisions: how much to invest (i.e. the rate of investment) and which projects to invest in. These decisions are interdependent, and both are taken, in the simplest version of the theory, with the same objective in view, namely, maximization of the firm's value to its owners. Undertaking investments alters the investment opportunities, not just by removing some, but also by creating others. The firm is always selecting the best opportunities it is aware of, but undertaking investments makes it aware of new opportunities since it learns by experience and by investing. Some of its investment consists of the cost of searching for opportunities. This is the case for research and development expenditures, which must be widely interpreted to include market research, investigation of new sources of supply, hire of business consultants and efficiency experts, etc. But almost any investment expenditure, by changing the world, enables something new to be learned. There is therefore no reason why investment opportunities should become exhausted. At any one time, it is true, the firm will be aware of a limited number of opportunities, some better than others. Hence increasing the rate of investment can be expected to worsen its average quality: more means worse. But there is no reason to suppose that the set of opportunities gets worse (or better) on average as time passes. Of course, it may do so, and for a particular firm there are bound to be fluctuations in the average quality of opportunities with which it is confronted. However, for the whole economy, and also for the *representative firm* in it, these fluctuations will tend to cancel out, and it is reasonable to assume as a working hypothesis that the average quality is unchanged. That, indeed, is the simplest hypothesis to take if we want to explain the stylized facts of growth.

It is, all the same, this hypothesis which is the most novel part of the theory, and which is therefore most in need of explanation and defence, and to this I now turn.

The inexhaustible cake of investment opportunities

Let us begin by considering a whole closed economy, the world if you like, as if it were a single enterprise. I do this so as to exclude learning from outside, and to include all the learning that there is. Later, I consider individual enterprises which learn a great deal from outside themselves. Let us also pose the problem in terms of the theory just described. In the formal development of the theory, I envisage the set of investment opportunities confronting a

representative firm at a given time as a sort of cake. The dimensions of the cake refer to the changes in output and employment resulting from investment, but I will not trouble you with the precise magnitudes. All I ask you to imagine now is that there is this cake of investment opportunities in front of the firm, out of which it is going to cut for itself a slice. The question is, what happens when this slice is taken? There are three possibilities which have been suggested at various times:

- (a) There is what might be called the stagnationist view, which received some support from Keynes,¹ and which may also have been the view of some classical economists. According to this view, investment opportunities are slowly exhausted, so that the returns to investment gradually fall. The slices taken out of the cake are then not replaced, and succeeding slices cut nearer and nearer to the point at which rates of return have fallen to zero, which could be reached, according to Keynes, within a generation. This view would make sense if we were omniscient, since then we would know the best way forward, and we would choose the highest yielding investments first, and they would be succeeded by investments yielding less and less. In such a world there would be no invention or discovery. It is clear that we are not in that world, and that the view just described does not accord with experience or the stylized facts of growth.
- (b) There is the orthodox view, according to which investment opportunities are being created at an exogenous rate by 'technical progress'. If the slices match this rate, each slice will be the same as the previous one, since fresh cake will be made just as fast as it is being eaten away. The stylized facts of growth can then be explained. If the slices are too big, the cake's growth cannot keep up with them and returns to investment will fall, and, if the slices are too small, returns to investment will rise. I have already criticized this view. It fails to explain why the cake grows. If that is attributed to the advance of science, why do we not double the rather small amounts invested in that way and reap the enormous returns? Is it plausible to suppose that a failure to invest for a 100 years

¹ 'On such assumptions [i.e. full employment and a "not disproportionate" rate of investment] I should guess that a properly run community equipped with modern technical resources, of which the population is not increasing rapidly, ought to be able to bring down the marginal efficiency of capital in equilibrium approximately to zero within a single generation.' (Keynes (1936), p. 220.)

would be followed by a long period when returns to investment would be very high? The view has been defended on a priori grounds, and the only empirical evidence in support of it that I know of is the importance of the unexplained residual in many growth accounting studies. However, these studies have all measured the contribution of investment incorrectly, as I have already pointed out. There is no residual to be explained.

- (c) There is the view I am defending that investment recreates investment opportunities. That the opportunities recreated should leave the same cake there as before is a working hypothesis which has the merit that it is consistent with the stylized facts of growth. That is an empirical justification, and I intend to provide further empirical justification, although not in this lecture, by using the model to explain a variety of phenomena. Perhaps that is all that needs to be said. However, some a priori reasons can also be advanced in defence. Investment leads to change and change is essential to learning. The more and the faster the change, the more and the faster the learning. Invention is a form of investment. There is no evidence that technical 'exhaustion' sets in, and the rate of invention is maintained if the rate of investment is maintained. At first sight it may seem puzzling that good opportunities should be replaced by good and bad by bad, that being the implication that the cake is unchanged. On reflection, however, that seems reasonable. One way in which undertaking investment creates further investment opportunities is by imitation, and so undertaking a very productive investment leads to very productive imitation, while a bad investment sets a bad example. More generally, the scope for learning from very productive investments should be greater, one might think, than the scope for learning from mistakes—although, like any generalization in economics, there will be exceptions. There is a temptation to think that investment opportunities which are *not* taken remain available, so that if good ones are not taken now they can be taken later, for example. But this temptation must be resisted. Investment opportunities do not remain in the cake unchanged until they are taken. In general, all are changed by the investments that *are* undertaken, so that what might have been good if done earlier could cease to be worthwhile—or could become even better.

So much for a whole, closed, economy. Let us now consider an

individual firm's investment opportunities. To some extent it will learn from its own investments and, in so far as that is the case, the preceding argument applies. To a great extent, however, firms learn from other firms' investments. There is then a flow of new ideas which is exogenous to the firm, and it might seem that the orthodox view (b) must then apply.

I certainly do not wish to deny that there *is* an exogenous flow of new ideas which partly determines a firm's investment opportunities. But what is often forgotten is that, along with *this* flow, which is favourable inasmuch as it continually widens those opportunities, there is an opposite flow which is continually narrowing them. This adverse flow consists of the attempts being made by competitors of the firm to capture its markets and bid away its labour or materials. Both the favourable and the adverse flows result from the investments of other firms.

I believe the net effect of these flows is favourable to a typical individual firm in the following sense. The faster is the rate of investment by other firms, the better will be the set of investment opportunities confronting the typical firm. In my model, this shows up as an enlargement of the cake. This effect seems most obvious if one thinks of R & D investment by others (including investment in universities and research institutes). But one should not think of the effect as being *confined* to R & D investment. *Any* investment by others may suggest investment opportunities to the typical firm, although some more than others. I am not here considering the effect of faster investment in permitting a faster growth of *demand*. That is an important additional effect, which arises because of market imperfections. I have discussed the role of demand elsewhere.¹ The effects being discussed here are due to learning and competition.

Although faster investment elsewhere increases and improves the set of opportunities for the typical firm, I do not think that it makes them *grow*, or grow faster. For a given rate of investment elsewhere the set for the typical firm is given. The alternative view would be that a greater rate of investment elsewhere makes the set grow faster for the typical firm. The investment opportunities would accumulate faster so that if our firm were to maintain its former rate of investment it would find its marginal rate of return continually increasing. I do not find that plausible, essentially because of competition from other firms. If the individual firm does not take up the better opportunities becoming available when investment elsewhere rises, then others will. The simplest

¹ See Sargent and Scott (1986).

plausible assumption is, then, that the set increases as investment elsewhere increases, but for given investment elsewhere the set of investment opportunities remains essentially unchanged.

I do not mean by this that they are unchanged in a *physical* sense. Because of changes elsewhere in the economy, a firm's investment opportunities will be continually different in common-sense terms. New inventions, new products, new capital goods, new markets, and new people will all alter the actual things which the firm must do to exploit its opportunities. The only sense in which I assume that the firm's investment opportunities are unchanged is in *economic* terms. Specifically, it is the case of investment opportunities which is constant: the relations between incremental output, incremental employment and investment. That it *is* unchanged is, of course, at best an approximation, and one which I expect to hold best for the typical or average firm.

Two qualifications must be made to the assumption of a constant set of investment opportunities. First, it does seem possible for parts of the world economy to lag behind other parts, and to benefit, as a result, from 'catch-up'. This phenomenon has been investigated by others. What may happen in such cases is that both the favourable stream of new inventions etc. and the adverse stream of competitive pressures are shut off from a particular country for various reasons. Subsequently, the country is opened up to such influences. Investors in that country can then reap the benefits of access to the new inventions etc. Since competition will not have forced up wage-rates, investment opportunities will on balance be enhanced, although the lack of investment in the past will limit this enhancement.¹ As investment proceeds, and the country catches up, wage-rates will be bid up and eventually the set of opportunities will become normal. This process of catching up can be envisaged for a country, or even possibly for a region within a large country, but hardly for an individual *firm*. The latter must compete in the same market as other firms for labour, and so is unlikely ever to be in a situation in which wage-rates are much out of line with those paid by its competitors.

The second qualification is that the assumption that the set of investment opportunities is constant in the sense described is only a working first assumption. No one can foretell the future, and it

¹ In developing countries, the lack of investment in infrastructure, and the low level of education of the work force, are often cited as factors which limit the profitability of private investment. This provides a counter-example to the conventional assumption of diminishing returns to the capital stock.

is entirely possible that the set will shrink or expand. It is also possible that, in the past, there have been periods in which the set has shrunk or expanded. I know of no law of nature which requires a constant set. What has happened to the set in the past is a matter for empirical investigation. Some have claimed to be able to detect long waves of invention which should show up as expansions and contractions in the set of opportunities. My own investigations suggest a movement in one direction only, happily that of expansion. However that may be, I suggest that for the *individual firm* the best working assumption at any given time will be that the set will remain constant.

I have merely sought to show that an inexhaustible, constant, set of investment opportunities is a reasonable working hypothesis. The utility of the hypothesis can be demonstrated only by using it to explain growth phenomena. For the present, it is to be hoped that you are at least ready to open your minds to the possibility that this is one example where you can both have your cake and eat it.

Conclusion

Orthodox theories of economic growth make its main determinants non-economic. Usher's¹ dictum 'no technical change, no growth' (meaning no labour productivity growth) sums up one of the theories' most striking conclusions, which appears to be supported by many empirical studies. I have tried to show where the empirical studies have gone wrong, and why a better dictum is 'no investment, no growth'. If I am right, economics has more to say about the causes of growth, and that should please economists. The determinants of the volume and efficiency of investment are restored to the centre of attention, and both have long been the concern of economists. We can, with renewed confidence in their importance, study the behaviour of firms, project appraisal, the working of capital markets, the determinants of saving, systems of taxation and their impact on savings, and the volume and pattern of investment, all of which are very relevant to economic growth in the long run. We do not have to abdicate to scientists and engineers, or even to those economists who specialize in the study of technical change. I do not deny for a moment the interest and importance of their work, but I hope I have convinced you that economists should be able to contribute much to an explanation of economic growth.

¹ See Usher (1980), ch. 12.

REFERENCES

- ABRAMOVITZ, M. (1956), Resource and Output Trends in the United States since 1870, *American Economic Association Papers* (May), pp. 5-23.
- BOSKIN, M. J. (ed.) (1979), *Economics and Human Welfare* (New York: Academic Press).
- CARRÉ, J.-J., DUBOIS, P., and MALINVAUD, E. (1976), *French Economic Growth* (London: Oxford University Press).
- COLLARD, D. A., HELM, D. R., SCOTT, M. FG., and SEN, A. K. (eds.) (1984), *Economic Theory and Hicksian Themes* (Oxford, Clarendon Press).
- Conference on Research in Income and Wealth (1951): *Studies in Income and Wealth*, xiv (New York: National Bureau of Economic Research).
- DENISON, E. F. (1962), *The Sources of Economic Growth in the United States and the Alternatives Before Us*, Supplementary Paper No. 13 (New York: Committee for Economic Development).
- assisted by POULLIER, J.-P. (1967), *Why Growth Rates Differ* (Washington D.C.: The Brookings Institution).
- (1974), *Accounting for United States Economic Growth 1929-1969* (Washington D.C.: The Brookings Institution).
- and CHUNG, W. K. (1976), *How Japan's Economy Grew So Fast* (Washington D.C.: The Brookings Institution).
- (1979), *Accounting for Slower Economic Growth* (Washington D.C.: The Brookings Institution).
- (1985), *Trends in American Economic Growth 1929-1982* (Washington D.C.: The Brookings Institution).
- FABRICANT, S. (1954), Economic Progress and Economic Change, *34th Annual Report of the National Bureau of Economic Research* (New York).
- HACCHE, G. (1979), *The Theory of Economic Growth* (London: Macmillan).
- JORGENSEN, D. W., and GRILICHES, Z. (1967), The Explanation of Productivity Change, *Review of Economic Studies*, xxxiv. 249-83.
- KALDOR, N. (1957), A Model of Economic Growth, *Economic Journal*, lxxvii. 591-624.
- KENNEDY, C., and THIRLWALL, A. P. (1972), Technical Progress: a Survey, *ibid.*, lxxxii. 11-72.
- KEYNES, J. M. (1936), *The General Theory of Employment, Interest and Money* (London: Macmillan).
- KUZNETS, S. (1961), *Capital in the American Economy* (National Bureau of Economic Research: Princeton University Press).
- LUTZ, F. A., and HAGUE, D. C. (eds.) (1961), *The Theory of Capital* (Proceedings of a Conference held by the International Economic Association, London: Macmillan).
- MATTHEWS, R. C. O., FEINSTEIN, C. H., and ODLING-SMEE, J. C. (1982), *British Economic Growth 1856-1973* (Oxford: Clarendon Press).
- OHKAWA, K., and ROSOVSKY, H. (1973), *Japanese Economic Growth* (London: Oxford University Press).
- ROBINSON, J. (1962), *Essays in the Theory of Economic Growth* (London: Macmillan).
- SARGENT, J. R., and SCOTT, M. FG. (1986), Investment and the Tax System in the U.K., *Midland Bank Review* (Spring), pp. 5-13.
- SCOTT, M. FG. (1976), Investment and Growth, *Oxford Economic Papers*, xxviii. 317-63.

- SCOTT, M. FG. (1981), The Contribution of Investment to Growth, *Scottish Journal of Political Economy*, xxviii. 211-26.
- SOLOW, R. M. (1957), Technical Change and the Aggregate Production Function, *Review of Economics and Statistics*, xxxix. 312-20.
- (1970), *Growth Theory* (Oxford: Clarendon Press).
- United States Department of Commerce (1972), The Measurement of Productivity, *Survey of Current Business*, vol. lii, no. 5, part II.
- USHER, D. (1980), *The Measurement of Economic Growth* (Oxford: Basil Blackwell).