Robert M. Solow - Prize Lecture

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1987
Robert M. Solow

Lecture to the memory of Alfred Nobel, December 8, 1987

Growth Theory and After

I have been told that everybody has dreams, but that some people habitually forget them even before they wake up. That seems to be what happens to me. So I do not know if I have ever dreamt about giving this Lecture. I know that I have been in this room before, but that was in real life, and I was awake. If I have given this lecture in my dreams, there is no doubt that the topic was the theory of economic growth. I am told that the subject of the lecture should be "on or associated with the work for which the Prize was awarded." That is pretty unambiguous. But I would not even wish to use the leeway offered by the phrase "associated with." Growth theory is exactly what I want to talk about: for itself, for its achievements, for the gaps that remain to be filled, and also as a vehicle for some thoughts about the nature of theoretical research in macroeconomics, and empirical research as well.

Growth theory did not begin with my articles of 1956 and 1957, and it certainly did not end there. Maybe it began with *The Wealth of Nations*; and probably even Adam Smith had predecessors. More to the point, in the 1950s I was following a trail that had been marked out by Roy Harrod and by Evsey Domar, and also by Arthur Lewis in a slightly different context. Actually I was trying to track down and relieve a certain discomfort that I felt with their work. I shall try to explain what I mean in a few words.

Harrod and Domar seemed to be answering a straightforward question: when is an economy capable of steady growth at a constant rate? They arrived by noticeably different routes, at a classically simple answer: the national saving rate (the fraction of income saved) has to be equal to the product of the capital-output ratio and the rate of growth of the (effective) labor force. Then and only then could the economy keep its stock of plant and equipment in balance with its supply of labor, so that steady growth could go on without the appearance of labor shortage on one side or labor surplus and growing unemployment on the other side. They were right about that general conclusion.

Discomfort arose because they worked this out on the assumption that all three of the key ingredients - the saving rate, the rate of growth of the labor force, and the capital-output ratio - were given constants, facts of nature. The saving rate was a fact about preferences; the growth rate of labor supply was a demographic-sociological fact; the capital-output ratio was a technological fact.

All of them were understood to be capable of changing from time to time, but sporadically and more or less independently. In that case, however, the possibility of steady growth

would be a miraculous stroke of luck. Most economies, most of the time, would have no equilibrium growth path. The history of capitalist economies should be an alternation of long periods of worsening unemployment and long periods of worsening labor shortage.

The theory actually suggested something even more dramatic. Harrod's writings, especially, were full of incompletely worked out claims that steady growth was in any case a very unstable sort of equilibrium: any little departure from it would be magnified indefinitely by a process that seemed to depend mainly on vague generalizations about entrepreneurial behavior. You may remember that John Hicks's *Trade Cycle* book, which was based on Harrod's growth model, needed to invoke a full employment ceiling to generate downturns and a zero-gross-investment floor to generate upturns. Otherwise the model economy would have run away.

Keep in mind that Harrod's first *Essay* was published in 1939 and Domar's first article in 1946. Growth theory, like much else in macroeconomics, was a product of the depression of the 1930s and of the war that finally ended it. So was I. Nevertheless it seemed to me that the story told by these models felt wrong. An expedition from Mars arriving on Earth having read this literature would have expected to find only the wreckage of a capitalism that had shaken itself to pieces long ago. Economic history was indeed a record of fluctuations as well as of growth, but most business cycles seemed to be self-limiting. Sustained, though disturbed, growth was not a rarity.

There was another implication of the Harrod-Domar model that seemed unsound. If the condition for steady growth is that the savings rate equal the product of the growth rate of employment and a technologically-determined capital-output ratio, then a recipe for doubling the rate of growth in a labor surplus economy was simply to double the savings rate, perhaps through the public budget. Well, not *simply*: we all knew then - as I am not sure we all know now - that doubling the *ex ante* saving rate would not double the *ex post* saving rate unless something were taking care of the *ex ante* investment rate at the same time. (I hope these strange Latin phrases are still understood in Stockholm in 1987!) In underdeveloped countries, however, where the appetite for new capital is likely to be pretty strong, the recipe looked usable. I believe I remember that writings on economic development often asserted that the key to a transition from slow growth to fast growth was a sustained rise in the savings rate. The recipe sounded implausible to me. I can no longer remember exactly why, but it did.

That was the spirit in which I began tinkering with the theory of economic growth, trying to improve on the Harrod-Domar model. I can not tell you why I thought first about replacing the constant capital-output (and labor-output) ratio by a richer and more realistic representation of the technology. I know that even as a student I was drawn to the theory of production rather than to the formally almost identical theory of consumer choice. It seemed more down to earth. I know that it occurred to me very early, as a natural-born macroeconomist, that even if technology itself is not so very flexible for each single good at a given time, aggregate factor-intensity must be much more variable because the economy can choose to focus on capital-intensive or labor-intensive or land-intensive goods. Anyway, I found something interesting right away.

It would sound silly for me to explain in any detail to this audience what I found. Nearly everyone who spends any time in this room already knows. The "neoclassical model of economic growth" started a small industry. It stimulated hundreds of theoretical and empirical articles by other economists. It very quickly found its way into textbooks and into the fund of common knowledge of the profession. Indeed that is what allows me to think that I am a respectable person to be giving this lecture today. Nevertheless I must summarize the outcome in a couple of sentences, so that I can move on to the more interesting questions about what is still unknown or uncertain and remains to be found out.

Just allowing for a reasonable degree of technological flexibility accomplished two things. In the first place, the mere existence of a feasible path of steady growth turned out not to be a singular event. A range of steady states is possible, and the range may even be quite wide if the range of aggregative factor-intensities is wide. There are other ways in which an economy can adapt to the Harrod-Domar condition, but it still seems to me that variation in capital-intensity is probably the most important.

Secondly, it turned out to be an implication of diminishing returns that the equilibrium rate of growth is not only not proportional to the saving (investment) rate, but is independent of the saving (investment) rate. A developing economy that succeeds in permanently increasing its saving (investment) rate will have a higher level of output than if it had not done so, and must therefore grow faster for a while. But it will not achieve a permanently higher rate of growth of output. More precisely: the permanent rate of growth of output per unit of labor input is independent of the saving (investment) rate and depends entirely on the rate of technological progress in the broadest sense.

There was a third result that seemed useful and certainly helped to make the model appealing to economists. Earlier growth theory was mechanical or physical, not in any bad sense but in the sense that it was almost entirely a description of flows and stocks of goods. In the neoclassical model it was quite natural and practical to describe *equilibrium* paths and to work out the price and interest rate dynamics that would support an equilibrium path. It did not occur to me the time that in doing this I was bringing good news and bad news. The good news was that economists instinctively like to think that way, and the connection would help to get my professional colleagues interested in growth theory. Moreover, it is a good (that is, fruitful) instinct, whether one is dealing with a capitalist or a socialist economy. The bad news is that the connection is a bit too pretty and too interesting and unleashes a standing temptation to sound like Dr. Pangloss, a very clever Dr. Pangloss. I think that tendency has won out in recent years, as I shall try to explain later on, though it may be too late for me to pretend to be Candide.

When I look back now at the articles I wrote in the 1950s and 1960s on this general subject, I am struck and even a little surprised at how much effort went into broadening the technological framework of growth theory. I wanted to make sure that the model could accommodate the likelihood that new technology can only be introduced with the use of newly designed and produced capital equipment, that factor proportions might be variable only at the instant of gross investment and not after capital equipment had taken some particular form, and that enough flexibility could be achieved with discrete activities, even with only one activity so long as the length of life of capital goods could be chosen economically. And in every case I wanted to show that the appropriate commodity-price-

factor-price relations could be worked out and made intelligible in terms of the inherited instincts of economists. (In my case I had inherited them mainly from Knut Wicksell and Paul Samuelson.)

There were reasons for this special orientation, reasons that seemed pretty compelling at the time. In the first place, it was the introduction of some technological flexibility that had opened up growth theory to a wider variety of real-world facts and to a closer connection with general economic theory. It seemed important to make sure that these gains were not tied too closely to an indefensibly simple version of factor substitution. Secondly, I had already begun to do some empirical work making use of an aggregate production function with apparently meaningful and clearly surprising results. I was very skeptical about this device myself, and I knew that others would have doubts of their own. It seemed like a good idea to make sure that the method was capable, at least in principle, of dealing with the first few doses of realism. And, thirdly, I was already trapped in the famous "Cambridge controversy". I use the word "trapped" because that whole episode now seems to me to have been a waste of time, a playing-out of ideological games in the language of analytical economics. At the time I thought - and the literature gave some reason to think - that part of the argument was about marginalism, about smooth marginalism. So I wanted to be able to show that the conclusions of the theory and of its empirical implementation were not bound to that very special formulation. I guess it was worth doing, but it certainly did not pacify anyone.

There was one bad by-product of this focus on the description of technology. I think I paid too little attention to the problems of effective demand. To put it differently: a theory of equilibrium growth badly needed - and still needs - a theory of deviations from the equilibrium growth path. I can honestly say that I realized the need at the time. There is a brief section at the end of my 1956 article that deals in a perfunctory way with the implications of real-wage rigidity and with the possibility of a liquidity trap. That was just a lick and a promise. There was also a paragraph that I am prouder of: it made the point that growth theory provides a framework within which one can seriously discuss macroeconomic policies that not only achieve and maintain full employment but also make a deliberate choice between current consumption and current investment, and therefore between current consumption and future consumption. Only a few years later I had the memorable experience in the Kennedy-Heller Council of Economic Advisers of seeing those ideas written into the 1962 *Economic Report* (which is about to be republished by the MIT Press). The history of the past seven years in the United States suggests that the lesson has not yet been learned in Washington.

The problem of combining long-run and short-run macroeconomics has still not been solved. I will come back to it later on. This is the place for me to confess to (and explain away) a certain youthful confusion. In the early discussions of Harrod-Domar growth theory there was much talk about the intrinsic instability of equilibrium growth. "Instability" could and did mean two different things, and the meanings were not always clearly distinguished. It could mean that well-behaved equilibrium paths are surrounded by badly-behaved equilibrium paths, so that a small sideward step could lead to eventual disaster. Or it could

mean that instability applies to disequilibrium behavior, so that an economy that once strays from equilibrium growth would not automatically find its way back to *any* equilibrium growth path.

The original Harrod-Domar model seemed to be subject to both these difficulties. I think I showed that extension of the model took the sting out of the first sort of instability. The second sort, however, really does involve the integration of short-run and long-run macroeconomics, of growth theory and business-cycle theory. Harrod and many contemporary commentators went at this problem by making very special (and unconvincing) assumptions about investment behavior. I may not have been as clear then as I am now about the distinction between the two notions of instability. Today I would put the unsolved problem as follows. One of the achievements of growth theory was to relate equilibrium growth to asset pricing under tranquil conditions. The hard part of disequilibrium growth is that we do not have-and it may be impossible to have-a really good theory of asset valuation under turbulent conditions. (1987 is an excellent year in which to make that observation!)

One important tendency in contemporary macroeconomic theory evades this problem in an elegant but (to me) ultimately implausible way. The idea is to imagine that the economy is populated by a single immortal consumer, or a number of identical immortal consumers. The immortality itself is not a problem: each consumer could be replaced by a dynasty, each member of which treats her successors as extensions of herself. But no short-sightedness can be allowed. This consumer does not obey any simple short-run saving function, nor even a stylized Modigliani life-cycle rule of thumb. Instead she, or the dynasty, is supposed to solve an infinite-time utility-maximization problem. That strikes me as far-fetched, but not so awful that one would not want to know where the assumption leads.

The next step is harder to swallow in conjunction with the first. For this consumer every firm is just a transparent instrumentality, an intermediary, a device for carrying out intertemporal optimization subject only to technological constraints and initial endowments. Thus any kind of market failure is ruled out from the beginning, by assumption. There are no strategic complementarities, no coordination failures, no prisoners' dilemmas.

The end result is a construction in which the whole economy is assumed to be solving a Ramsey optimal-growth problem through time, disturbed only by stationary stochastic shocks to tastes and technology. To these the economy adapts optimally. Inseparable from this habit of thought is the automatic presumption that observed paths are equilibrium paths. So we are asked to regard the construction I have just described as a model of the actual capitalist world. What we used to call business cycles - or at least booms and recessions are now to be interpreted as optimal blips in optimal paths in response to random fluctuations in productivity and the desire for leisure.

I find none of this convincing. The markets for goods and for labor look to me like imperfect pieces of social machinery with important institutional peculiarities. They do not seem to behave at all like transparent and frictionless mechanisms for converting the consumption and leisure desires of households into production and employment decisions. I can not imagine shocks to taste and technology large enough on a quarterly or annual time scale to be responsible for the ups and downs of the business cycle. But now I have to report

something disconcerting. I can refer you to an able, civilized and completely serious example of this approach and suggest that you will find it very hard to refute. You can find non-trivial objections to important steps in the argument, but that would be true of any powerful macroeconomic model.

There is a dilemma here. When I say that Prescott's story is hard to refute, it does not follow that his case can be proved. Quite the contrary: there are other models, inconsistent with his, that are just as hard to refute, maybe harder. The conclusion must be that historical time series do not provide a critical experiment. This is where a chemist would move into the laboratory, to design and conduct just such an experiment. That option is not available to economists. My tentative resolution of the dilemma is that we have no choice but to take seriously our own direct observations of the way economic institutions work. There will, of course, be arguments about the *modus operandi* of different institutions, but there is no reason why they should not be intelligible, orderly, fact-bound arguments. This sort of methodological opportunism can be uncomfortable and unsettling; but at least it should be able to protect us from foolishness.

Since what I have just said goes against the spirit of the times, I would like to be very explicit. No one could be against time-series econometrics. When we need estimates of parameters, for prediction or policy analysis, there is no good alternative to the specification and estimation of a model. To leave it at that, however, to believe as many American economists do that empirical economics begins and ends with time series analysis, is to ignore a lot of valuable information that can not be put into so convenient a form. I include the sort of information that is encapsulated in the qualitative inferences made by expert observers, as well as direct knowledge of the functioning of economic institutions. Skepticism is always in order, of course. Insiders are sometimes the slaves of silly ideas. But we are not so well off for evidence that we can afford to ignore everything but time series of prices and quantities.

After this methodological digression, I should remind you of the direction of my main argument. Growth theory was invented to provide a systematic way to talk about and to compare equilibrium paths for the economy. In that task it succeeded reasonably well. In doing so, however, it failed to come to grips adequately with an equally important and interesting problem: the right way to deal with deviations from equilibrium growth. One possible solution strikes me as wrong-headed: that is to deny the existence of an analytical problem by claiming that "economic fluctuations" are not deviations from equilibrium growth at all, but examples of equilibrium growth. My impression is that belief in this story is more or less confined to North America. Maybe the experiences of European economies does not lend itself to this interpretation at all. What alternatives are there?

It will not do simply to superimpose your favorite model of the business cycle on an equilibrium growth path. That might do for very small deviations, more in the nature of minor slightly autocorrelated "errors." But if one looks at substantial more-than-quarterly departures from equilibrium growth, as suggested for instance by the history of the large European economies since 1979, it is impossible to believe that the equilibrium growth path itself is unaffected by the short- to medium-run experience. In particular the amount and directions of capital formation is bound to be affected by the business cycle, whether through gross investment in new equipment or through the accelerated scrapping of old

equipment. I am also inclined to believe that the segmentation of the labor market by occupation, industry and region, with varying amounts of unemployment from one segment to another, will also react back on the equilibrium path. So a simultaneous analysis of trend and fluctuations really does involve an integration of long-run and short-run, or equilibrium and disequilibrium.

The simplest strategy is a familiar one from other contexts. In a completely aggregated growth model the relevant prices are the real wage and real rate of interest. Suppose they are both rigid, or merely adjust very slowly to excess supplies in the markets for labor and goods. (The more usual assumption is that only the wage is sticky; but in Wicksell's own native habitat we should allow for a divergence between the "natural" and "market" rates of interest.) Then the economy may be away from any full equilibrium path for a long time. During that time its evolution will be governed by a short-run dynamics much like everyday business-cycle theory.

The most interesting case to consider is one where real wage and rate of interest are stuck at levels that lead to excess supply of labor and goods (saving greater than investment ex ante). This is the sort of configuration we have come to call "Keynesian." The big difference is that net investment may be positive or negative; industrial capacity may be rising or falling. The economy may eventually return to an equilibrium path, perhaps because "prices are flexible in the long run" as we keep telling ourselves. If and when it does, it will not return to the continuation of the equilibrium path it was on before it slipped off. The new equilibrium path will depend on the amount of capital accumulation that has taken place during the period of disequilibrium, and probably also on the amount of unemployment, especially long-term unemployment, that has been experienced. Even the level of technology may be different, if technological change is endogenous rather than arbitrary.

This is the sort of amendment that I mentioned in 1956, but did not pursue very far. There is now an excellent exploratory sketch by Edmond Malinvaud using this fix-price approach to growth theory. As you would expect, an important role is played by the investment function. When I referred earlier on to the difficult problem of asset valuation away from an equilibrium path, this is what I meant. We are reduced to some more or less plausible formulation guided by more or less robust econometric results and by whatever we think we know about investment decision-making in real firms. Malinvaud emphasizes "profitability" as a determinant of investment, but he also emphasizes that the precise meaning of profitability is unclear whenever the future is unclear.

The main result of Malinvaud's analysis is a clarification of the condition under which a "Keynesian" steady state is possible, and when it is locally stable, i. e. when it will be approached by an economy disturbed from a nearby equilibrium path. The unstable case is just as interesting, because it suggests the possibility of small causes having big results. All these stability arguments have to be tentative because the interest rate and real wage are assumed to be fixed while quantities move. That is not an adequate reason to dismiss the results in a purist spirit; but obviously the research program is not complete.

A sketch by Malinvaud is as good as a book by someone else. My own inclination - it is just an inclination - is to try a slightly different slant. Thinking about the ambiguousness of the concept of profitability and its relation to investment reminds one that many firms react to

changed circumstances precisely by changing their prices. The obvious alternative to a model with sticky prices is a model with imperfectly competitive price-setting firms. Then, of course, one can no longer speak in any simple way of excess supply of goods. But we can find something just as interesting; the possibility of many coexisting equilibrium paths, some of which are unambiguously better than others. (Usually the better ones have higher output and employment than the worse ones, so something like recession makes an appearance anyway.) The interaction of growth and business cycle can then take a slightly different form: alternation of good and bad equilibria is not just a simple averaging.)

This sort of model is now pretty familiar in a static context, where it can make good working sense of the notion of "effective demand." Firms will naturally condition their actions on beliefs about economic aggregates. Frank Hahn and I are working on extending it to a model of overlapping generations, so that it would be easy to convert any stationary equilibrium state into a growing steady state. Preliminary indications are that the thing can be done. There is a hope, therefore, that either the fix-price approach or the imperfect-competition approach can allow us to talk sensibly about macroeconomic policy in a growth context.

In my 1956 paper there was already a brief indication of the way neutral technological progress could be incorporated into a model of equilibrium growth. It was a necessary addition because otherwise the only steady states of the model would have constant income per person and that could hardly be a valid picture of industrial capitalism. Technological progress, very broadly defined to include improvements in the human factor, was necessary to allow long-run growth in real wages and the standard of living. Since an aggregate production function was already part of the model, it was natural to think of estimating it from long-run time series for a real economy. That plus a few standard parameters - like saving rate and population growth - would make the model operational.

Estimating an aggregate production function was hardly a new idea, but I did have a new wrinkle in mind: to use observed factor prices as indicators of current marginal productivities, so that each observation would give me not only an approximate point on the production function but also an approximate indication of its slopes. I am pretty sure that this idea was suggested to me by equilibrium growth theory. I want to emphasize that I did not then have any notion I was doing something intensely controversial.

The first few paragraphs of my 1957 article are thoroughly ambivalent, not about the method but about the use of aggregate data on inputs and output. After expressing my doubts I went ahead in a pragmatic spirit. One can not do macroeconomics without aggregative relationships; and at least for the moment there is no substitute for macroeconomics. The only way I can account for the intensity of controversy over this point is to ascribe it to the belief that there is something intrinsically ideological about the notion that profits on "capital" represents the return to a factor of production as imputed by the market. John Bates Clark may have thought, a century ago, that distribution according to marginal products was "just" but no modern economist, no modern "bourgeois" economist, would accept that reasoning.

Anyway, the main result of that 1957 exercise was startling. Gross output per hour of work in the U. S. economy doubled between 1909 and 1949; and some seven-eighths of that increase could be attributed to "technical change in the broadest sense" and only the remaining eight could be attributed to conventional increase in capital intensity. Actually Solomon Fabricant at the National Bureau of Economic Research had come up with a similar breakdown for a slightly earlier period, using methods with less in the way of analytic foundation. I think I had expected to find a larger role for straightforward capital formation than I actually found; I will come back to that point soon.

The broad conclusion has held up surprisingly well in the thirty years since then during which time "growth accounting" has been refined quite a lot, especially by Edward Denison. The main refinement has been to unpack "technical progress in the broadest sense" into a number of constituents of which various human-capital variables and "technological change in the narrow sense" are the most important. To give you an idea of the current state of play I shall quote Denison's most recent estimates for the United States.

Taking the period from 1929 to 1982 and smoothing away the business cycle, he finds that the real non-residential business output increased at an average rate of 3.1 percent a year. The problem now is to parcel this out among a number of basic determinants of growth. Denison estimates that a quarter of it can be attributed to increased labor input of constant educational level. Another 16 percent (i.e. about 1/2 percent a year) is credited to the increased educational qualifications of the average worker. The growth of "capital" accounts for 12 percent of the growth of output; this is coincidentally almost exactly what I found for 1909-1949 using my original method, of which Denison's is in some ways a practical refinement. Then Denison imputes 11 percent of total growth to "improved allocation of resources" (by which he means such things as the movement of labor from low-productivity agriculture to higher productivity industry). Another 11 percent goes to "economies of scale" (but this must be a very insecure imputation). Finally 34 % of recorded growth is credited to "the growth of knowledge" or technological progress in the narrow sense. If you add up these percentages, you will see that Denison has accounted for 109 percent of measured growth. Miscellaneous factors must then have reduced the growth of output by nine percent of 3.1 percent, or just under 0.3 percent a year. (These negative factors could include such things as investment in environmental improvement, which uses resources but does not appear in measured output, though it may of course be very valuable.)

This detailed accounting is an improvement on my first attempt, but it leads to roughly the same conclusion. Remember that I distinguished only three factors: straight labor, straight capital, and residual "technical change". Denison decomposes the residual into five components, but the flavor is very similar.

The similarity is brought out more strongly if one looks at Denison's results on a "per person employed" basis. Real output per person employed grew by 1.7 percent per year between 1929 and 1982. Labor input per person employed accounted for - 23 percent of this. That sounds strange; but means mostly that hours worked per year per person employed fell during the period, so that the average employed person provided less straight labor time. I will not go over the full imputation. All I want to point out is that education per worker accounts for 30 percent of the increase in output per worker and the advance of knowledge

accounts for 64 percent in Denison's figures. Thus technology remains the dominant engine of growth, with human capital investment in second place. One does not have to believe in the accuracy of these numbers; the message they transmit is pretty clear anyway.

That is meant as a serious remark. If I may revert to methodological propaganda again, I would like to remind my colleagues and their readers that every piece of empirical economics rests on a substructure of background assumptions that are probably not quite true. For instance, these total-factor-productivity calculations require not only that market prices can serve as a rough and ready approximation of marginal products, but that aggregation does not hopelessly distort these relationships. Under those circumstances, robustness should be the supreme econometric virtue; and over-interpretation is the endemic econometric vice. So I would be happy if you were to accept that the results I have been quoting point to a qualitative truth and give perhaps some guide to orders of magnitude. To ask for much more than that is to ask for trouble. I would also like to quote the profound warning issued by the leading student of the statistics of baseball - it hangs in my office - "No amount of (apparent) statistical evidence will make a statement invulnerable to common sense".

The mention of common sense brings to mind another aspect of this story, still unsettled in the literature. In the beginning, I was quite surprised at the relatively minor part the model ascribed to capital formation. Even when this was confirmed by Denison and others, the result seemed contrary to common sense. The fact that the steady-state rate of growth is independent of the investment quota was easy to understand; it only required thinking through the theory. It was harder to feel comfortable with the conclusion that even in the shorter run increased investment would do very little for transitory growth. The transition to a higher equilibrium growth path seemed to offer very little leverage for policy aimed at promoting investment.

The formal model omitted one mechanism whose absence would clearly bias the predictions against investment. That is what I called "embodiment", the fact that much technological progress, maybe most of it, could find its way into actual production only with the use of new and different capital equipment. Therefore the effectiveness of innovation in increasing output would be paced by the rate of gross investment. A policy to increase investment would thus lead not only to higher capital intensity, which might not matter much, but also to a faster transfer of new technology into actual production, which would. Steady-state growth would not be affected, but intermediate-run transitions would, and those should be observable.

That idea seemed to correspond to common sense, and it still does. By 1958 I was able to produce a model that allowed for the embodiment effect. A certain amount of simplicity was lost, because the stock of capital could no longer be regarded as a homogenous lump. One had to keep track of its age structure; but that was precisely the point. Anyhow the model was workable even if it was not neat. If common sense was right, the embodiment model should have fit the facts significantly better than the earlier one. But it did not. Denison, whose judgment I respect, came to the conclusion that there was no explanatory value in the embodiment idea. I do not know if that finding should be described as a paradox, but it was at least a puzzle.

In the course of preparing this lecture, I came across a recent working paper by Professor Edward N. Wolff (of New York University) which offers a longer-run perspective on this matter. Wolff compiled data for seven large countries (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States) covering the whole century from 1880 to 1979. He also paid special attention to the postwar period 1950-79. These particular countries were selected for data availability only, so they can not be considered a representative sample. Wolff's result is therefore only suggestive, but it is an interesting suggestion.

For each of the countries he calculates the average growth rate of Total Factor Productivity (i.e., what I have called the rate of technical progress in the broad sense) and also various measures of the speed of investment. (For instance he looks at the growth rate of the capital stock, the growth rate of the capital-labor ratio, and the average investment quota itself.) Then, looking across countries, he finds a very strong positive correlation between the rate of technical progress and the speed of investment. His interpretation is that this provides strong confirmation of the embodiment hypothesis: if we suppose that all these countries had access to roughly the same pool of technological innovations, then it appears that the ones that invested fastest were best able to take advantage of the available knowledge. That is certainly one reasonable interpretation and it is one I like. Keep in mind that, by using total factor productivity, Wolff has already "given" to investment its traditional function of increasing productivity by increasing capital intensity, so the remaining correlation is between investment and the *shift* of the aggregate production function.

To be faithful to my own methodological precepts, however, I should remind you that other interpretations are also possible. For example, it could be the case that some countries are better able to exploit the common pool of technological progress than others, for reasons that have nothing to do with the rate of capital formation; but in exactly those technologically progressive countries investment is most profitable, so naturally the rate of investment is higher. Or else rapid technical progress *and* high investment could both be the result of some third factor, like the presence of conditions that encourage entrepreneurial activity. High investment and fast technical progress will then go together.

I can not argue strongly one way or the other. But at least the way remains open for a reasonable person to believe that the stimulation of investment will favor faster intermediate-run growth through its effect on the transfer of technology from laboratory to factory.

Before I finish, perhaps I should point out that it is possible to combine most of the building-blocks I have been discussing in a small but fairly complete econometric model. If that were not possible, I would find the ideas less interesting. It has in fact been done. One example is the "annual growth model of the U.S. economy" due to Bert Hickman and Robert Coen.

This is a model whose production side is completely aggregated and is, in fact, just exactly the sort of thing I have been talking about. (The demand side is disaggregated, but that is not important now.) The full equilibrium paths of the Hickman-Coen model are exactly those made familiar by growth theory, a little more general because the determination of saving and the evolution of the labor force are looked after in more detail.

That part is quite straightforward. In some recent exercises, however, Hickman and Coen have started a serious study of deviations from equilibrium growth in exactly the spirit recommended by Malinvaud and by me. They allow for real wage rigidity, and they model their producing sector as a price-setting monopolistic competitor. Now investment does not have to be equal to full-employment saving, except in full equilibrium. Periods of boom and stagnation can appear, and do appear, to almost no one's surprise. There can be "Keynesian" and "classical" unemployment. Indeed there can be both at the same time: the real wage might be too high to allow full employment with existing capital stock, while at the same time aggregate demand is inadequate to take off the market what firms would wish to produce. Changes in the real wage could have demand-side and supply-side effects.

All this sounds very good, sounds just like the macroeconomics that pragmatic Americans and Swedes have practiced all along. I can not vouch for the Hickman numbers, but they are at least sensible. They show, by the way, that high-real-wage induced unemployment was negligible in the U.S. between 1959 and 1978, and was then again dwarfed by low-demand induced unemployment in 1981 and 1982. I do not know what their story is for the years after 1982, but the fact that I would like to know speaks well for the model.

In this brief review of the goals and achievements of growth theory I have referred as much to the work of others as to my own. That is more than mere modesty: the choice reflects my belief that any successful line of economic analysis is almost certain to be a group product. We attach names to ideas for good and bad reasons, but useful ideas are usually worked out and critically refined by a research community. I have some faith that the ideas of "neoclassical" growth theory are viable just because they have attracted a research community, even a rather diverse community: Lucas and Prescott build on the basic model, and so do Malinvaud and "sunspot" theorists like Karl Shell and others.

When I read Robert Frost's lines from "The Black Cottage":

> Most of the change we think we see in life
> is due to truths being in and out of favor

it occurred to me at once that they sound altogether too much like economics. Some of that feeling is inevitable, and not necessarily to be regretted. The permanent substructure of applicable economics can not be too very large because social institutions and social norms evolve, and the characteristics of economic behavior will surely evolve with them. I believe also that part of the changeability of economic ideas on a shorter time-scale is our own doing. It comes from trying too hard, pushing too far, asking ever more refined questions of limited data, over-fitting our models and over-interpreting the results. This, too, is probably inevitable and not especially to be regretted. You never know if you have gone as far as you can until you try to go further.

Naturally I hope that growth theory can serve in both ways: as a background on which to hang multi-sector models that probably try to do more that can be done, and as a framework for simple, strong, loosely quantitative propositions about cause and effect in macroeconomics. For both roles, it appears to me, the fundamental intellectual need is for a common understanding of medium-run departures from equilibrium growth. That is the

stuff of everyday macroeconomics. It has been going on in English-speaking countries since Keynes and in Sweden since Lindahl and the Stockholm School. It is going on in both places today.

References

Denison, E., *Trends in American Economic Growth, 1929-1982*, Brookings Institution, 1985
Hickman, B., "Real Wages, Aggregate Demand, and Unemployment," *European Economic Review* December 1987.
Hickman, B. and R. Coen, *An Annual Growth Model of the U.S. Economy*, North-Holland, 1976
Malinvaud, E., "Notes on Growth Theory with Imperfectly Flexible Prices," in *Modern Macroeconomic Theory*, (ed. J-P Pitoussi) Basil Blackwell, 1983
Prescott, E., "Theory Ahead of Business Cycle Measurement," Working Paper, Federal Reserve Bank of Minneapolis, February 1986
Solow, R., "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, February 1956
- "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, August 1957
- "Investment and Technical Progress" in *Mathematical Methods in the Social Sciences, 1959* (ed. K. Arrow, S. Karlin, P. Suppes) Stanford University Press, 1960
- J. Tobin, C. von Weizsaecker, and M. Yaari, "Neoclassical Growth with Fixed Factor Proportions," *Review of Economic Studies*, April 1966
Wolff, E., "Capital Formation and Long-Term Productivity Growth," Working Paper, C. V. Starr Center for Applied Economics, New York University, September 1987

From *Nobel Lectures, Economics 1981-1990*, Editor Karl-Göran Mäler, World Scientific Publishing Co., Singapore, 1992

Addendum, August 2001

My Nobel Lecture in 1987 sketched the intellectual environment in which the neoclassical model of economic growth was first worked out in the 1950s, and it then discussed some of the paths along which growth theory had evolved afterwards. I took it for granted that growth theory should be an integral part of pragmatic macroeconomics, a model to be applied to our own economies. My view was, and still is, that the most urgent current analytical need was for a way of fitting together short-run macroeconomics, when the main action consists of variations in aggregate demand, with the long run factors represented by the neoclassical growth model, when the main action is on the supply side. Another way of saying this is that short-run and long-run models of macroeconomic behavior need a way to merge in a practical macroeconomics of the medium run.

Not much has been accomplished in that direction since 1987. This may be because of the profession's enthusiasm for Real Business Cycle Theory, an enthusiasm I did not share (and I explained why, briefly, in the lecture). Nevertheless there are some recent signs of progress. A few members of the RBC school have realized that empirical success will not come unless at least a few realistically modelled imperfections and frictions are absorbed into the model. But some of these are just the imperfections and frictions - "sticky prices" for

instance - that are the true microunderpinnings of more "Keynesian" macroeconomic models. At the same time, some advocates of a more "Keynesian" approach have adopted some of the favorite practices of the RBC school, for instance a more explicit use of dynamic general equilibrium models. Creating a seamless, unified macroeconomics of the medium run may be a difficult or impossible task because the reality is just too messy to be described by any small number of basic principles. But a start has been made.

In retrospect, the most surprising aspect of the 1987 lecture must be that it does not even mention the idea of endogenous technological progress. This was primarily a matter of timing. The first of Paul Romer's series of papers was published only in 1986, and Robert Lucas's paper did not appear in print until 1988 (although I had undoubtedly read it in manuscript before 1987). But of course I had no way of knowing that "endogenous growth theory" would develop into a veritable mass movement during the 1990s. I had not forgotten my old friend Kenneth Arrow's 1962 paper on Learning by Doing. But I already knew that, interesting as that model was, it made no real difference to the basic propositions of the neoclassical growth model. I later elaborated on this point in two Arrow Lectures published as *Learning from Learning by Doing*." Endogenous growth theory as it developed was a much more drastic change.

We know now that endogenous growth theory led to an avalanche of papers that has recently slowed a little, but only a little. It is easy to see why the idea was so popular. The models offered the possibility of having a theory of the steady-state growth rate itself, instead of treating it as an exogenously given, if sometimes changing, fact of life. But there was an even more important attraction, I think. The nature of the theory was such that one could easily find feasible, even fairly traditional, policies that would influence the long-term growth rate. Adding a couple of tenths of a percentage point to the growth rate is an achievement that eventually dwarfs in welfare significance any of the standard goals of economic policy. Who would not be excited?

It is important to understand how it was that endogenous growth theory could offer this prospect. The earliest models simply assumed constant returns to *capital*, or to the set of factors of production that can be accumulated, like capital. (They were called "AK models" for this reason.) Thus, for example, if output is just proportional to capital and saving-and-investment is proportional to output, then investment is proportional to capital and the saving?investment rate enters the factor of proportionality. So a higher saving-investment rate means a higher ratio of investment to capital, i.e. a higher growth rate of capital, and therefore a higher growth rate of output. But we think we know policies that will increase the *level* of saving-investment from given output. This easy passage from influencing a level to influencing a growth rate is what makes the theory so powerful. The trouble is that constant returns to capital is a highly special, pinpoint assumption. This is one of those cases where "approximately" will not do. But exactly constant returns to capital is not very plausible empirically, and has no convincing theoretical foundation either.

Deeper and more interesting models soon emerged in the endogenous growth tradition. Some of them focus on the creation and accumulation of human capital, others on the process of technological invention and innovation (and the temporary monopolies that go with it). There is also a flourishing group of "Schumpeterian" models that emphasize the rivalry (or occasional complementarity) between an innovation and its predecessors. It

seems to me that work along these lines might eventually tell us something interesting and useful about the role of knowledge in the economy and society. We have always realized that there is an important endogenous element in the development of new technology; all those businesses investing millions in research are not suffering from a mass delusion.

Up to now, however, I have found that even these deeper and more circumstantial models of endogenous growth all rest at some key point on an essentially arbitrary linearity assumption, on the claim that the rate of growth of this is a function of the level of that, where "that" is some fairly simple and accessible variable that can be maneuvered by policy. Of course such a claim can be true, but the ones I have seen have been neither empirically verifled nor overwhelmingly plausible *a priori*.

One possible interim conclusion is that the literature strains too singlemindedly for a theory of the steady-state growth rate, as if exponential growth were an intrinsic characteristic of the time-paths of industrial economies, and not merely a convenient approximation to a segment of such a time-path. An alternative strategy might be to begin with unprejudiced empirical study of the determinants of the speed of technological innovation or the accumulation of human capital or the evolution of quality ladders for consumer goods. These things are already studied by economists (and other social scientists) independently of their relevance for growth theory. Undoubtedly any valid generalizations would have to be simplified for inclusion in an aggregative growth model, but that is only to be expected. I suppose it is always possible that some of those powerful linearity assumptions will prove to be valid, but I think it unlikely. One of the advantages of cheap computation and easy numerical integration is that the short-cut of steady state analysis is no longer needed; one can calculate directly the path traced out by a plausibly enhanced model of growth, and how it responds to the standard policy moves.

Another, much less prominent, line of thought may be worth mentioning. It goes back to the 1950s when Nicholas Kaldor tried to produce a coherent growth model based entirely on relationships among rates of growth, conspicuously without any explicit function relating inputs and outputs. A similar idea was taken up by Maurice Scott (*A New View of Economic Growth*, 1989), with much more attention to arguing the realistic case for such a relationship. (See also T. van de Klundert's *Growth Theory in Historical Perspective*, 2001), especially Chapter 7.)

I have never been able to convince myself that these convenient assumptions make sense, except as they are derived from more primitive input-output relations. The examples in the literature do not seem to have this property. My own attempt to get at the same underlying intuition (also dating from the 1950s) was to study "vintage models" in which technological innovations can only be effective when they are "embodied" in new capital goods. This attempt also languished, although it has recently been revived in a series of papers by Greenwood, Hercovitz, Krusell and Wolff. It seems possible that the proper place for such models to come into their own is in the elusive macroeconomics of the medium run.

<http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1987/solow-lecture.html>