REFLECTIONS ON GROWTH THEORY

ROBERT M. SOLOW

Abstract

This note contains some general and idiosyncratic reflections on the current state of neoclassical growth theory. It expresses some surprise at the lack of attention both to multi-sector growth models and to multi-country models with trade and capital flows. It also suggests that there might be value in further analysis of some old topics like capital–labor substitution with an expanded definition of capital, and the interaction of growth and medium-run phenomena (or, to put it differently, the interaction of demand-side and supply-side variations).

Keywords

I cannot remember what words Charles Dickens put in the mouth of The Ghost of Christmas Past. This is not the Cratchit family dinner anyway, and there is no one to play the role of Scrooge. But there is no doubt that I am here in roughly the capacity of the Ghost. So I will make up something for him to say.

We are nearing the 50th anniversary of the neoclassical model of growth; astonishingly, it is still alive and well. There is not really any competing model. In the broad sense in which I use the term, the "endogenous growth" models of Romer and Lucas and their many successors are entirely neoclassical. So the *basic* model has survived for 50 years. I emphasize "basic" because progress, in theory and in practical analysis, has come mainly from extending the basic model at the edges. The territory of growth theory has expanded to include more topics in what used to be border areas. This is not necessarily exactly the same thing as "endogenizing" these borderline topics. There is a lot in the *Handbook* about the influence of background forces like "institutions" on the evolution of technology or total factor productivity. Some of it is in the mood of the "New Growth Theory" but not all of it. Much of it just wants to be explicit about background forces without trying to absorb them into the model.

I will come back to some of the extensions of growth theory; but it is also interesting to contemplate a few of the territories into which the theory has not expanded. For example, I suspect that early on one would have expected much more work on multi-sector growth models than there has been. Not that there has been none: Leif Johansen had an early book, oriented toward planning. Luigi Pasinetti has written extensively on the sorts of structural changes to be expected along a trajectory, arising from such inevitable factors as differing income elasticities of demand for different goods. In a very different vein, there was a whole literature stemming from the von Neumann model, which now seems to have gone out of favor. Xavier Sala-i-Martin's chapter in the *Handbook* reviews some worthwhile developments and promises others.

In the early stages there was active exploration of two-sector models, culminating in the book by Duncan Foley and Miguel Sidrauski, but it petered out fairly soon. The reason was probably internal-intellectual rather than any feeling that the applications were unimportant. The usual, perfectly reasonable, choice was to distinguish between a consumer-good-producing sector and an investment-good sector. (Agriculture and Industry was another possible split, but mainly in the development context.) I have the feeling that too much in those models turned out to depend on differences in factor-intensity between the sectors. We have very little in the way of facts or intuition about that issue, and there was no reason to expect or postulate any systematic pattern that could lead to exciting results.

It is also a little odd that there was not more in the way of open-economy growth modelling. There was of course the well-known book by Gene Grossman and Elhanan Helpman; it attracted attention more for its analysis of endogenous technology and quality ladders than for trade and capital flows. Jaume Ventura's chapter in the *Handbook* records the current state of play. I can only say that 40 years ago I would have been expecting to see more research in these areas than actually turned up.

These undercultivated subfields would have had, could still have, important practical applications. I will just mention three examples.

(a) The creation and enlargement of the European Union have been modelled in several places; but I do not think that we have had the insights that could come from embedding this question in a formal multi-sector, multi-country growth model. (How much does it matter where the R&D is done? Are there potential gains from more migration within the EU? And so on.)

(b) Is there anything deeper to be learned from the fact that, after a couple of decades of catch-up to the U.S. in productivity and TFP, the large European economies seem recently to have stagnated relatively or even fallen back a bit?

(c) If the U.S. (and the EU) were to impose their own local environmental standards on their poor-country trading partners, what could be expected to happen to factor prices, real income levels and growth rates in the poor countries? All these things have been thought about, of course, but something might have been gained had appropriate growth models been easily available.

A slightly larger and slightly different question has to do with China: How should growth theory be applied to such a large, diverse, almost dual economy, especially in a world of rapidly increasing trade and international investment? Certainly the many-industry, many-country aspects must matter, but sheer geographical size and regional diversity may require special treatment. Chinese economists have already started applying modern growth theory to their problems; but some departures from the usual might be in order.

While the Ghost is going on about might-have-beens, I will allow him a couple of paragraphs on a topic that he has grumbled about before. Neoclassical growth theory is about the evolution of potential output. In other words, the model takes it for granted that aggregate output is limited on the supply side, not by shortages (or excesses) of effective demand. Short-run macroeconomics, on the other hand, is mostly about the gap between potential and actual output. (There is an important modern school of macroeconomics that assumes this distinction away, and makes the growth model explain short-run fluctuations too. It would be a digression to discuss that issue here.) On the older view – this is after all the Ghost talking – some sort of endogenous knitting-together of the fluctuations and growth contexts is needed, and not only for the sake of neatness: the short run and its uncertainties affect the long run through the volume of investment and research expenditure, for instance, and the growth forces in the economy probably influence the frequency and amplitude of short-run fluctuations. This terrain is sometimes described as the economics of the medium run.

It too has been undercultivated by growth theory. It has not been entirely ignored; but I have the impression that growth theorists simply write this off as a trivial perturbation that can not be allowed to deflect their own preoccupation with steady-state growth. For example, the work of Robert Coen and Bert Hickman, who actually do try to embed a serious demand side in a serious growth-model framework, and implement the result econometrically, is generally ignored by growth theory. There must be other scattered forays in this direction; I have taken a casual step or two myself. It should be a more

central part of growth theory proper. To put it differently, it would be a good thing if there were a unified macroeconomics capable of dealing with trend and fluctuations, rather than a short-run branch and a long-run branch operating under quite different rules. My quarrel with the real business cycle and related schools is not about that; it is about whether they have chosen an utterly implausible unifying device.

I can now turn from the things that growth theory has not accomplished to the things that it has done, in particular the way it has expanded outside the confines of a narrow model. The main effort has quite properly gone into the endogenization of changes in technology (or more broadly TFP, though usually with technology in mind) and changes in the stock of human capital. In both cases the popular early theoretical models had features that I personally found unappealing and, in policy terms, misleading for reasons that I have pursued elsewhere, and do not intend to repeat now. On the whole, better ideas have driven out worse ones as they are supposed to do. Both lines of research – technology and human capital – have led to a welcome emphasis on social norms and institutions as enabling or limiting factors or even as actual sources of growth.

The extent of interest in such ideas is represented explicitly in the *Handbook* by the chapters by Acemoglu–Johnson–Robinson, Greif, Alesina, Parente–Prescott, and implicitly by others. This emphasis on the role of institutions at least opens up the possibility – about which I am now more optimistic than I once was – of connecting up growth theory with the problem of economic development, in which issues of institutional change are clearly central. My own prejudice – Ghosts are allowed, even encouraged, to have prejudices – is that there may have been a premature tendency to assimilate growth and development, abetted by the vogue for cross-country regressions. A country is a country, one might say, just another point in $(n + 1)$-dimensional space, although loud squeals from the data have sometimes forced a restriction of the sample to OECD countries. This is something that needs to be straightened out; and detailed analysis of institutions is probably a better method than cross-country regressions.

The breathtakingly broad sweep of the story-line proposed by Daron Acemoglu and colleagues is irresistibly fascinating. Much of it has the ring of truth. I must confess nevertheless to a certain skepticism about firm conclusions at this level of generality, especially when they bear on "ultimate" causality. "Good" political institutions can certainly make the path to growth-friendly economic institutions shorter and smoother. But there are cases of "bad" – autocratic – governments opting for enforceable property rights and other "good" economic institutions, possibly in the belief that economic success will ultimately strengthen the hand of the autocrats themselves. Singapore and early post-war South Korea are examples; one of them evolved toward political democracy and the other did not.

The interaction of political institutions and the available stock of human capital can be very complex. Very poor countries are usually characterized by very bad governments and very deficient human capital, and these probably reinforce one another. I would not find it hard to accept the notion that there is no universally reliable way to escape this trap. Some countries succeed and others fail, for reasons that may be totally obscure *ex ante* and only partially and tentatively explicable *ex post*.

Whatever generalizations we are prepared to accept, however, there has to be a next stage: after history has made it plain that secure property rights and markets are better for growth than mere hierarchical rent-extraction, what do you do for an encore? The devil is in the details. What sort of patent protection provides the best mixture of incentives for innovation and diffusion? How should the free-rider problem intrinsic to non-firm-specific training best be handled? Do alternative feasible norms for corporate governance have any significant implications for growth?

These and many other institutional choices are practically invisible on the Acemoglu scale, but they bulk pretty large if you are considering alternative policies for a growing capitalist economy or for a transition economy. A very similar point is made by Rodrik in his chapter of the *Handbook*, but in a slightly different context. It is reassuring that many of the same considerations that preoccupy Acemoglu et al. also figure at the nitty-gritty level: those institutional choices have real distributional consequences that can in turn help or hurt the vested interests that in turn may be able to block those very choices. Can the protagonists be bought off or overcome politically or satisfied by compromise? The familiar elasticities and marginal whatnots come back into play as soon as one tries to face up to those questions in systematic but practical ways.

The translation of any "institutional" question into the language of an aggregative model is always tricky. The concepts and quantities that appear in an economic model need not be capable of expressing what a knowledgeable observer would like to say about institutional differences. An example of this occurs in the pleasing and informative chapter by Philippe Aghion and Peter Howitt. It is an incidental matter that caught my eye because it relates to some independent work of my own. The standard Schumpeterian wisdom is that active competition is bad for innovation because it erodes entrepreneurial rents too soon. But the standard empirical–historical finding is that active competition is associated with productive innovation. Aghion and Howitt find a way to enlarge their model so that the competition-innovation nexus can in principle go either way. (It is a clever and useful device that the interested reader should study.) The trouble is that the competition that Aghion and Howitt can conveniently model is between an innovating monopolist and the competitive fringe still stuck with the older technology. But the typical way in which the absence of competition deadens economic performance is that regional or national monopolies are protected by legal or other barriers against competition from best-practice firms, or just better-practice firms, domestic or foreign. They can survive without innovating or adopting best practice and they do so. To incorporate that chain of causation in a standard growth model might be difficult. Is it worth doing? We will not know until somebody tries.

No one would claim that we now have a really good causal account of either technological and organizational progress or of the accumulation of human capital in all of its various forms. (Schooling is not the same thing as human capital.) As the *Handbook* shows, a lot of ingenuity has gone into that research and lot of ground has been covered; if there is a lot more ground still to be explored, that is hardly surprising. Nor is there any guarantee that everything one would want to know about these processes

is knowable. Some regression residuals represent not omitted variables but mere sound and fury, signifying nothing.

It is possible that the obvious importance and interest of these matters, combined with a less worthy temptation, the sheer convenience of the Cobb–Douglas production function, may have diverted attention from an older-fashioned topic, namely the substitution of (physical and human) capital for raw labor. In the beginning, one of the surprising implications of the neoclassical growth model was not merely that the steady-state growth rate was independent of the saving-investment rate, but perhaps even more that the (moving) equilibrium *level* of output per person apparently responds very weakly to changes in the saving-investment quota. To be more precise, the elasticity of output per worker with respect to the saving-investment rate is the ratio of the capital share to the labor share (in steady-state equilibrium). Back then the conventional numbers were 1/4 and 3/4, giving an elasticity of 1/3. The message appeared to be that as big an increase in the rate of investment as policy could manage would yield only a disappointing increase in productivity. The conventional numbers have changed, partly as a matter of fact and partly because the capital concept has been enlarged. Apart from that, there may be some interesting points of principle to be investigated. (I have to admit that I dwell on this because Olivier de La Grandville and I have written a substantial paper on the subject.)

The point I want to make here starts from the fact that the relative shares of capital and labor once seemed to be trendless in modern industrial economies; and that may still be a fair description. (This generally accepted fact provided a respectable justification for the addiction of theorists to Cobb–Douglas.) Nowadays, however, there is some basis for thinking that the capital share may have risen. It is useful in this connection to keep in mind that the aggregative elasticity of substitution, taken as measuring the responsiveness of relative factor prices to relative factor intensity, is not a purely technological concept (as John Hicks realized when he invented it in 1932). One quantitatively important way in which the aggregate economy can substitute capital for labor is through a factor-price-and-commodity-price-induced shift in demand from labor-intensive to capital-intensive goods and services. That sort of easy substitution, just like easy technical substitution, is also a way of fending off diminishing returns.

So it is at least thinkable that the aggregative elasticity of substitution might be or might become fairly large, especially at high capital–labor ratios. It was part of the original neoclassical growth model that a large enough elasticity of substitution allows sustained growth in output per worker and capital per worker even without technological progress (and faster still with it).

Now, to get back to where I started, since the elasticity of substitution is larger than one in such a process, the (competitive) share of capital would be rising along with the capital–labor ratio, indeed rising toward one. But then the elasticity of productivity with respect to the saving-investment rate would be getting very large. Remember: I am not prepared to tell this story, only to suggest that the mechanics of the aggregative elasticity of substitution might be as interesting an object of study for growth theorists as the last

little twist on endogenous technical progress. There may be several ways to postpone or hold off the influence of diminishing returns to broadly-defined capital.

Discussion of these various facets of growth theory calls to mind a background decision that is rarely explicit: What sort of characteristic time interval are we talking about? Is growth theory 5-year, 20-year or 100-year economics? I know what I had in mind at the beginning: growth models were about intervals of time just long enough so that deviations above or below potential output would be small relative to the increment to output from beginning to end. An older terminology would have said that the model is about trend with business cycle removed; but that is awkward because the model itself is supposed to determine the trend, and I have doubts about the utility of the "cycle" idea. Anyhow that was then.

The range of time perspectives implicit in current research, as represented in the *Handbook*, is very broad. At one extreme, many macroeconomists propose to use a growth model to describe events quarter by quarter, presumably week by week if the data were available. At the other extreme, the focus on political economy and institutional change seems to call for marking the clock in intervals of a century more or less. I have already suggested that I think the first choice seems willfully to ignore or deny economic events that ought to be important in macroeconomics. The very long perspective tends to treat as mere perturbations or disturbances changes that ought to be at the center of growth theory itself. If I were using a growth model to interpret century-long time series, I would want to re-calibrate the parameters a couple of times in between if the data seemed to call for it.

I think I am driven back to the earlier convention. To be excessively concrete: an economy growing at an average rate of three percent a year will roughly double in 25 years. Episodes of recession or overheating amounting to five percent of total output would cover most short-run fluctuations, and would probably count as relatively small compared with overall growth. This suggests to me that the natural habitat of growth-theoretic explanations is time-spans of 25 to perhaps 40 or 50 years. Anything much shorter is likely to mix up supply-side and demand-side effects, and anything much longer runs the risk of overlooking some events that ought to be accounted for explicitly. That cannot be called an error; the very-long-run perspective has its own rewards. But then we needs some nomenclature to distinguish it from the 30-year perspective.

These scattered observations and remarks seem to coalesce – at least they do for me – in a general reflection. I have set it down before, obscurely, and it may bear repeating. Growth theory has focused mainly on the steady state. This is partly because accounting for Nicholas Kaldor's "stylized facts" was one of the early goals of growth modelling; and they were essentially a description of a steady state. In addition, it is much easier to work out general and robust properties of steady states than to do the same for transition paths. So the emphasis on steady states is both explicable and reasonable.

There is a side-effect, however, that strikes me as not so good. Somehow the convention has become established that a policy aimed at "growth" is by definition a policy that will increase the steady-state growth rate. A policy package that merely increases output by 10 percent at every point along an already established trajectory is somehow disap-

pointing. You might take this as just another example of the principle that where there is life there is hype. But for growth theorists it has led to a premium on models that do offer a direct connection between easily manageable policies and the steady-state growth rate. The early crop of so-called "AK" models was the predictable response. I thought they were uninteresting theory, in the sense that they more or less assumed what they purported to prove, and also misleading guides to policy, in that they made something look easy that is in fact very difficult. The fashion for such models seems to have waned. William Easterly's chapter in the *Handbook* demonstrates rather convincingly that these apparently easy levers on the growth rate are indeed illusory, as we should have known.

If you look closely, however, even more serious models of endogenous growth seem to depend, at a key point, on the blunt assumption that, for some important X, $dX/dt = A(\ldots)X$, where $A(\ldots)$ is a function of one or more easily manipulable level variables. Maybe that is a logical necessity: if you want an exponential solution, you better have a linear differential equation. One has to accept that piece of wisdom. The real point, however, is that any such linearity assumption, because it is so powerful, ought to require much more convincing justification than it gets in the standard models of endogenous technological change or accumulation of human capital.

I wonder if it would be a fair summary of the *status quo* to say that the broad neo-classical model of growth is widely accepted as a valid description of the mechanism of economic growth in advanced economies; most current research is aimed at unpacking, understanding and testing those aspects that the basic model takes as given. The Ghost could go along with that, subject to a couple of minor amendments already stated. The mechanism itself needs extension so that it can cover international flows of goods, capital and technology (and perhaps labor), so that it can better describe interactions between demand-driven fluctuations and the path of potential output, and so that it can allow explicitly for the existence of many goods and sectors with different technologies and different demand conditions. These needs were foreseeable and foreseen early in the story; they are merely unfinished old business.