COORDINATION ISSUES IN LONG-RUN GROWTH*

PETER HOWITT

Department of Economics, Brown University, Providence, RI

Contents

*For helpful suggestions I thank, without implicating, Rob Axtell, David Laidler, Richard Lipsey, Malcolm Rutherford, Leigh Tesfatsion and three anonymous referees.

Handbook of Computational Economics, Volume II: Agent-Based Computational Economics, edited by Ken Judd and Leigh Tesfatsion

Abstract

Economic growth depends not only on how people make decisions but also upon how their decisions are coordinated. Because of this, aggregate outcomes can diverge from individual intentions. I illustrate this with reference to the modern literature on economic growth, and also with reference to an older literature on the stability of full-employment equilibrium. Agent-based computational methods are ideally suited for studying the aspects of growth most affected by coordination issues.

Keywords

Growth, coordination, innovation, stability, agent-based systems

JEL Classification: E00, O33, O40

1. Introduction

Economic growth, like most economic phenomena, depends on the incentives that people face when making decisions. Measured in per-capita terms, growth cannot be sustained indefinitely unless some combination of capital, skills and knowledge grows without bound. So we cannot understand long-term growth without some understanding of what induces people to invest in capital, skills and knowledge. Reduced-form AK theories focus on the intertemporal choices by households that underlie capital accumulation. Innovation-based growth theories of the sort that Philippe Aghion and I have been working on for some time[1] focus on the R&D decisions of profit-seeking business firms that lead to the innovations that raise the stock of disembodied technological knowledge. Human-capital based theories focus on the time-allocation decisions of households investing in education and training. In all cases, changes that impinge on the incentives of the decision makers affect an economy's long-run growth rate.

Some writers have gone so far as to make incentives the *sine qua non* of growth economics. Thus Easterly (2001, p. 289) states that "Prosperity happens when all the players in the development game have the right incentives," and quotes approvingly (p. xii) from Steven Landsburg that "People respond to incentives; all the rest is commentary." To Lucas (2002, p.17) what matters above all is the incentives facing household decision makers:

> For income growth to occur in a society, a large fraction of people must experience changes in the possible lives they imagine for themselves and their children, and these new visions of possible futures must have enough force to lead them to change the way they behave, the number of children they have, and the hopes they invest in these children: the way they allocate their time.

My purpose in this essay is to take issue with this exclusive focus on incentives and the logic of choice. Not to deny that incentives matter for economic growth but to assert that much else matters also, and that much of what also matters is ideally suited for study by computational methods.

Economies are large complex systems that can be studied at different levels. Macroeconomic issues, which involve the functioning of the system as a whole, need to be studied at a coarser

[1] Aghion and Howitt (1998a)

level than microeconomic issues involving the behavior of just one market or just a small group of individuals, households or business firms. A clear understanding of the entire system would be obscured by focusing on a detailed analysis of these constituent working parts, just as a clear understanding of ocean tides would be obscured by focusing on the molecules of water in the ocean, or a clear view of a pointillist painting would be obscured by examining each dot one at a time. The system as a whole is not a macrocosm of its individual parts and the parts are not microcosms of the whole. Instead, as Schelling (1978) has argued forcefully, macro behavior can depart radically from what the individual units are trying to accomplish. So when you stand back the details become hard to see but patterns emerge that were not visible from up close.

Thus my primary objection to the agenda laid out by Lucas and others is that it is likely to involve a fallacy of composition. Incentives and decision-making are properties of the constituent parts of an economy, whereas economic growth is a property of the system as a whole. If the economy functioned as a macrocosm of its parts then focusing on incentives would yield a clear picture of the growth process. But I believe it is not. What matters at the macro level is not just how individual transactors formulate their plans but also the nature of their interactions with each other and with their environment. In short, an economy's growth performance often depends not so much on how people make their decisions as it does on how those decisions are coordinated, or in some cases how the decisions become uncoordinated.

One of the virtues of the ACE approach to economics, as outlined by Tesfatsion (2005), is that it forces one to make explicit the mechanisms through which individual actions are coordinated, for better or worse. That is, in order to make a model "dynamically complete," in Tesfatsion's terminology, one has to specify what will happen from any given set of initial conditions, including those in which different people are acting on the basis of inconsistent beliefs and hence in which aggregate outcomes will necessarily diverge from individual intentions. Another virtue of the ACE approach is that it provides a method for discovering a system's "emergent properties," i.e. those properties that are not inherent in the individual components. Thus it seems ideally suited for studying those aspects of the growth process that go beyond the Lucas agenda.[2]

[2] Work that has used the ACE approach for studying technological change, the ultimate mainspring of long-run growth, is surveyed by Dawid (2005).

2. The representative agent model and its limitations[3]

The idea that the economy as a whole can behave very differently from what the individual transactors are trying to accomplish is hardly original. Indeed one of the oldest themes of economic theory is that things are not as they seem to the individual. The classical economists delighted in pointing out how the unconstrained pursuit of maximal profit by competing sellers would end up minimizing their profit. Smith's attack on mercantilism was based on the idea that although the accumulation of precious metals would make an individual wealthy it would not do the same for a nation. Keynes argued that the unemployment rate was determined not by individual labor-supply decisions but by what was happening in product markets and in the financial system. The first textbooks promoting the Keynesian revolution highlighted the paradox of thrift, according to which the attempt by individual households to save more could end up reducing the economy's overall level of saving. One of Friedman's central arguments in promoting Monetarism was that people who favor a policy of cheap money don't realize that in the long run this will cause higher interest rates. Thus what happens to profits, wealth, unemployment, saving or interest rates depends not so much on individual choices and intentions as on how those choices and intentions are coordinated. Focusing on the incentives faced by individuals trying to influence the variable would produce the wrong answer. A broader perspective is needed.

But by the start of the 21st Century, the education of a macroeconomist no longer included any warnings against the fallacy of composition. On the contrary, the very foundations of modern macroeconomics, as practiced in academic research and taught to graduate students, is the belief that macro variables are best understood by focusing on the details of decision-making by individual households and firms. In such theories, macroeconomic variables such as interest rates, wage rates and unemployment rates reflect intertemporal substitution and time-allocation decisions on the part of a representative household, whose behavior is indeed a small-scale replica of the system as a whole. High unemployment reflects a disincentive to work, low saving a disincentive to abstain from current consumption, and high interest rates a high rate of individual time preference or a low elasticity of intertemporal substitution in consumption. The fallacy of division that this approach entails is just the dual of the fallacy of composition. In

[3] The limitations of the representative agent model have been examined extensively by Kirman (1992).

effect, these twin fallacies play an even bigger role in a macroeconomist's education than they did a generation ago; the difference is that instead of being taught as pitfalls to be avoided they are now presented as paradigms to be emulated.

How this transformation in economics took place is a long story that I cannot begin to unravel here. The transformation is clearly related to the rational-expectations revolution started by Lucas's justly celebrated 1972 *Journal of Economic Theory* paper, which provided a micro-foundation for a macro theory that claimed to reconcile the long-run neutrality of money with short-run non-neutrality. When rational expectations was adopted by the advocates of Keynesian economics as well as by its critics, the gap between micro and macro became not bridged but papered over. For the very idea that individual actions could have unforeseen consequences does not sit easily with the idea that everyone acts rationally, guided by an accurate model of how the overall economy works. Moreover, the very terminology of "rational" expectations draws one's attention to individual thought processes, obscuring the fact that the achievement of rational expectations is really a collective process requiring the coordination of what must initially have been non-rational expectations.

But clearly there is more to this transformation than rational expectations. The history of the development of Keynesian macroeconomics from the end of World War II was one of providing a choice-theoretic underpinning to the behavioral functions that comprise the IS-LM system. The representative household and firm played as much a part in this pre-rational-expectations theoretical development as they have since 1972. It seems that in seeking to provide a bridge between micro and macro, economists have been driven by a reductionist imperative to bring everything down to the level of individual choices and by an "irrational passion for dispassionate rationality."[4] Conventional acceptance of these attitudes makes it easy to dismiss as *ad hoc* or poorly grounded any theory that starts with behavioral rules not explicitly derived from rational foundations. Adherence to this standard makes it necessary to use something like the representative agent just to keep manageable a model of the whole economy that focuses sharply on the constituent parts. It also makes it necessary to assume away most of the coordination problems that would get in the way of rational expectations by blurring the link between individual choices and their consequences.

[4] The phrase, which I first heard from David Laidler, is commonly attributed to J.M. Clark.

To be sure, not all macroeconomists accept this representative-agent view of short-run macroeconomics, and much progress has been made recently in studying the coordination problems that might impede the formation of rational expectations (see for example Sargent, 1993 or Evans and Honkapohja, 2001). But there is still a widespread belief that the importance of coordination problems is limited to short-run theory, like the price-stickiness that can keep the economy away from its natural rate of unemployment in the short run or the informational imperfections that permit a short-run Phillips curve to be exploited by policy-makers. It is generally regarded as uncontroversial to model long-run phenomena like economic growth by assuming that aggregate variables are chosen rationally by some representative agent, whose incentives are therefore all that really matter for understanding the economy's performance.

Economics being an empirical science, the first question to ask of the agenda that Lucas and others have laid out is whether there is a prima facie case for believing that overall economic performance reflects the intentions of the individual decision makers. Is it really true that, to a first approximation, rich nations are those whose citizens have a lot of education, save a large fraction of their incomes and work long hours? More to the point, is it really true that nations that grow rapidly are those in which there is high investment in physical capital, education and R&D?

The evidence from the recent "development accounting" literature is not all that convincing. Although Mankiw, Romer and Weil (1992) tried to argue that 75 percent or more of the cross-country variation in per-capita GDP was accounted for by a simple Solow-Swan model in which the main variables were investment rates in physical capital and enrollment rates in education, the vast literature spawned by this provocative article has shown that these rates are themselves endogenous to income levels and also highly correlated with productivity. Thus it seems that countries with high investment and enrollment rates tend to be rich to a large extent just because they are also nations in which more output can be produced from any given amount of physical capital and education. Klenow and Rodríguez-Clare (1997) estimate that more than 60 percent of the cross-country variation of per-worker GDP is attributable to productivity rather than to the accumulation of physical and human capital.

When it comes to accounting for differences in growth rates, which is after all the primary objective of growth theory, the evidence for the incentive agenda is even less convincing. According to Klenow and Rodríguez-Clare, over 90 percent of the cross-country variation in growth rates of per-worker GDP is attributable to differences in productivity-growth rates rather than to differences in investment rates or enrollment rates. Thus it seems that almost everything to be explained by the theory lies in the Solow residual, which Abramowitz once called nothing more than a measure of our ignorance.

This is part of the evidence that inclines me towards innovation-based growth theory, since most of the effects of innovation work through productivity-growth rates. So is it really countries that spend a large proportion of their GDP on R&D that have the fastest productivity-growth rates? Coe and Helpman (1995) and Coe, Helpman and Hoffmaister (1997) have examined the cross-country relationships between growth rates and R&D intensities (the fraction of GDP spent on R&D) and found that there is indeed a powerful relationship, but what matters to an individual country is not so much its own R&D intensity as that of its major trading partners. This mirrors at the country level the result that one typically finds at the industry level (see for example Zachariadis, 2003). That is, the research efforts undertaken by firms in one country or one industry aimed at enhancing their own productivity end up enhancing productivity in other countries and industries. Presumably this reflects a process of technology spillover, or what is sometimes called "technology transfer." So here again, the behavior of a variable (one country's productivity or one industry's productivity) is an unintended consequence of the incentives faced at the individual level, a consequence that involves the channels through which individual transactors interact rather than the manner in which they decide to act.

3. Externalities and unintended side effects

As I have already observed, the professional consensus in macroeconomics seems to be that coordination issues are more important for short-run theory than for the theory of long-run growth. This is a legacy of the neoclassical synthesis, according to which sticky prices and informational imperfections are just transitory impediments to the smooth coordination of rational choices. More generally it reflects what Clower and I (1998) have called the "classical stability hypothesis," to the effect that in the long run the economy will converge to a coordinated state. Yet there are sound theoretical reasons for thinking that the process of economic growth brings with it a set of forces that widen the gap between individual intentions and aggregate outcomes rather than the reverse, and reasons for thinking that the growth process often exacerbates the impediments to smooth coordination rather than the reverse. The present section of the paper and the next elaborate on this point.

One reason why the growth process can widen the intention-output gap is the central role that externalities play in the process. The ultimate mainspring of growth is technological change, which is known to involve significant external effects: the empirical work on technology spillovers referred to above corroborates a plausible theoretical presumption that the ideas generated by R&D are hard to appropriate. Thus as Arrow (1969) argued, innovation tends to go under-rewarded because it confers much of its benefits on third parties. To complicate matters, Schumpeter's notion of creative destruction, which Aghion and I have developed in our work, involves a negative spillover that tends to give people too strong an incentive to perform R&D. That is, the firm performing R&D takes into account the prospective rents that would be created by a new product or process but does not take into account the rents that would be destroyed through obsolescence by the same innovation.

Externalities are hard to ignore in growth theory not just because of these substantive reasons but also because of the technical difficulties of coping with increasing returns to scale. That increasing returns is involved in one form or another once technology becomes endogenous has been recognized at least since Allyn Young (1928). In modern innovation-based theory increasing returns takes the form of a setup cost of research, which is independent of the size of the market to be served by the resulting innovations. Producing the first unit of a new product takes so much resource input for the original innovation and so much for the variable production cost. Producing each subsequent unit requires only the variable cost. Average cost is thus decreasing with the amount produced.

Indeed the upsurge of endogenous growth theory in the past two decades can arguably be attributed not so much to the new substantive ideas that it has produced as to the progress it has made in dealing with the technicalities of increasing returns. In particular, we know that a competitive equilibrium without externalities generally fails to exist in a world with ubiquitous decreasing cost. You need to introduce some combination of either pecuniary externalities (imperfect competition) or direct non-pecuniary externalities. What endogenous growth theory did was to borrow techniques for dealing with these externalities from other areas of economics (the Dixit-Stiglitz-Ethier model for dealing with imperfect competition and the concept of symmetric anonymous Nash equilibrium for dealing with non-pecuniary externalities) in order to develop manageable models of ideas that have been common among economic historians and specialists in the economics of technology for several generations.

How the growth theories that have been developed on these grounds can generate aggregate outcomes that contradict individual intentions is illustrated by a central result of Aghion, Harris, Howitt and Vickers (2001) concerning the effects of intellectual property protection on an economy's overall level of R&D and hence on its overall rate of technological progress. Weaker patent protection reduces the direct incentive for a firm in any given situation to perform R&D. Yet it can actually raise the aggregate level of R&D and hence raise the overall rate of technological progress. It does this through a "composition effect," which works as follows. Innovation takes place at the greatest rate in those industries where the leading firms are neck-and-neck; that is, where they produce using similar technologies. This is because profits are lowest in such industries and hence the incentive to escape competition by innovating is strongest. If patent laws were weakened, a firm with any given technological lead over its rivals would have its incentive to innovate blunted, but the steady-state distribution of lead sizes would also be changed; specifically, more firms would find themselves in the R&D-intensive situation of neck-and-neck competition because of a rival's success in imitating their technological capability. As a result, it can be shown theoretically that under a wide variety of circumstances there is a point up to which weaker patent laws will raise the economy's overall growth rate, even though the incentive for a firm in any given situation goes in the opposite direction.

Likewise, as Mokyr (1990) has argued, nations that experience the most rapid growth are not necessarily those in which people have the strongest incentive to develop new technologies but those which have developed the greatest tolerance for, and capacity to adjust to, the many negative side-effects of economic growth. Those negative side-effects are almost always the result of obsolescence – the destructive side of creative destruction. Because of obsolescence, technological change is a game with losers as well as winners. From the handloom weavers of early 19th century Britain to the former giants of mainframe computing in the late 20th century, many people's skills, capital equipment and technological knowledge have been devalued and rendered obsolete by the same inventions that have created fortunes for others. The conflict between winners and losers from new technologies is a recurrent theme in economic history, and the difficulty of mediating the conflict affects society's willingness to foster and tolerate economic growth.

Thus for example, ever since the introduction of machinery into manufacturing processes in the early part of the industrial revolution, people have been worried that economic growth could cause technological unemployment. Mainstream professional economists have tended to regard such popular concerns as fallacious, with a few notorious exceptions like Ricardo's (1821) chapter "On Machinery". The classical stability hypothesis leads one to believe that the unemployment created by any one technological innovation should be short-lived; those rendered unemployed will eventually find employment elsewhere. But this is not true if we look at an increase in the rate at which new technologies are being introduced rather than at a single innovation. As Aghion and I (1994) have argued, a faster pace of job-destroying innovations will raise the flow into unemployment in any given situation, and can thereby increase the steady-state (natural?) rate of unemployment.

Unemployment is more of a social problem in some countries than others. In the United States, for example, where wages are more flexible and employment regulations less restrictive, technologically induced unemployment is likely to be less of a social problem than in many European countries. But this just tends to exacerbate another common side-effect of rapid technological progress, namely rising wage-inequality. As many have pointed out, the last quarter of the 20th Century was a period of rapidly rising inequality, especially in the United States. Although public opinion often blames globalization for this rise in inequality, the culprit to which academic research points more often is skill-biased technological change. In short, the same phenomenon that caused high unemployment levels in Europe by destroying jobs seems to have caused high wage-inequality in the US by enriching those who can work with new technologies and driving those whose jobs are destroyed into less remunerative jobs.

To some extent this side effect is one that can be dealt with by more investment in education – by raising the number of people able to work profitably with new technologies instead of being displaced by new technologies. In principle this should help not just those whose skills are enhanced by more education but also those who remain relatively less skilled, whose wages should be lifted by their increasing relative scarcity. But recent theoretical research suggests at least two problems with this approach. One is that not all of the increase in inequality is explained by an increasing educational premium. Instead, roughly half of the overall increase is attributable to a rise in residual inequality, the inequality that is unexplained by education, experience or any other observable individual characteristic.

Aghion, Violante and I (2002) have argued that this is because whether or not someone is able to work with new technologies is often a matter of pure luck rather than of education levels. The luck factor is always there in the wage distribution; indeed we know that income inequality between identical twins tends to be about as large as within the whole population. But it was greatly leveraged by the IT revolution, not only because this was a general purpose technology that hastened the pace of technical change, and hence further raised the wages of those lucky enough to have just the right skills, but also because of the nature of IT. That is, because of the generality of computer technology and the associated reduction in communication costs, many of those lucky enough to be able to work on the leading edge of technology today have skills that can easily be marketed throughout the entire economy, rather than in just one sector, and they receive a compensation that is correspondingly enhanced. There is nothing that increased investment in human capital can do to counteract this particular side effect of economic growth.

The other problem that has been raised by theoretical research is the "market-size effect" that Acemoglu (2002) has explained. That is, because the cost of R&D takes the form of a setup cost, researchers tend to direct their efforts towards enhancing the productivity of factors that are relatively abundant in the economy rather than those that are relatively scarce; although the cost of either type of effort might be the same, the payoff is larger from enhancing a factor that is more widely used. Acemoglu shows how this can produce a positive feedback loop, whereby more education induces even more innovations that enhance the relative productivity of educated workers and hence increase their relative wage, which in turn induces even more people to become educated. This is just fine for those who are capable of joining in, but for the old and less able the situation is one of increasing relative poverty, one that would just be exacerbated by policies raising the incentive to acquire education.

Societies that are willing to cope with and possibly mitigate high unemployment and/or high inequality are thus likely to be those that put up the fewest impediments to the introduction and adoption of new technologies, and hence to be those that have the highest long-run growth rates. Of course incentives matter in this story, but not those that we would be led to examine by simple representative-agent models. What promotes growth in these stories is not the willingness of households to accumulate physical or human capital or the willingness of firms to engage in R&D but rather the willingness of politicians to permit side effects to persist or to devise institutions like unemployment insurance, redistributive schemes, relocation subsidies, etc., that alleviate the side effects. In short, economic growth is at least as much about mediating social conflict as it is about the individual virtues of thrift, study and hard work.

4. Uncertainty and the classical stability hypothesis

The main reason for thinking that the growth process can exacerbate coordination problems is the fundamental uncertainty of technological progress. Technological innovation is a destabilizing force that is constantly disrupting established patterns of economic activity, much like the disturbance term in a time-series process. But the path that technology follows is a highly non-stationary one which, while it may exhibit some aggregate patterns, is virtually unpredictable in its details. Thus from the point of view of the individual decision maker, an innovation is not something that simply alters the initial condition in some well understood dynamic game, but one that destroys the value of previous information and starts an adaptive learning process all over again. The more rapid the pace of innovation the more chaotic the process becomes, the less confidence people are able to place in history as a guide to the future, and therefore the more likely their individual plans are to be thwarted by unsuspected macro forces.

The unpredictability of technological progress is a major theme in the writing of Nathan Rosenberg, who has pointed out how technologies that were developed for one purpose very often had their major impact on something their discoverer was unaware of. Bell Labs, for example, where scientists invented the laser, was reluctant to patent it because in their opinion it had no conceivable industrial uses in the telephone industry (Rosenberg, 1994, p.223). Thomas Watson Sr., the founder of IBM, at first regarded the computer as a highly specialized scientific instrument with no potential commercial uses (Rosenberg, 1994, p.220). Technological developments in the sewing machine industry ended up having a major effect on automobiles, which had not yet been invented at the time of the discoveries (Rosenberg, 1963).

Writers like Brian Arthur (1989) have observed that because of this fundamental uncertainty, the pace and direction of innovation are necessarily guided by short-term considerations, even though they can lead society down irreversible paths whose long-run consequences are of great import, especially when there are "network externalities" involved. That is, the course of technological progress, rather than reflecting the intentions of those individuals that create it, is a social process driven by the largely unforeseen consequences of individual decisions. If these aggregate consequences are unforeseen at the level of the individuals involved then surely we have little chance ourselves of understanding them unless we look at them from a different level, presumably from the level of the system as a whole.

The disruptiveness of technological change is something that writers like Freeman and Perez (1988) and David (1990) have analyzed extensively. They argue that major technological changes come in waves, driven by what are now commonly called general purpose technologies (GPTs); that is, new technologies that are used throughout the economy, have a profound effect on the way economic life is organized, and give rise to a wave of complementary innovations. In the long run our standard of living has been greatly enhanced by the succession of GPTs introduced since before the first Industrial Revolution, including such things as the steam engine, electric power, and the computer.[5] However, the period during which a new GPT is being introduced can be a period of wrenching adjustment, not just at the level of the individual firm but for the economy as a whole.

There are many aspects to this adjustment cost that have been studied in the literature. Helpman and Trajtenberg (1998) emphasize the lost output that can occur because a GPT never arrives fully developed but instead requires the subsequent invention of a set of complementary components. During the period when the components are being developed, the new GPT will not yet be used to its full effect. Meanwhile the labor that is drawn into developing new components will be drawn out of producing final output. The result can be a fall in the overall level of output. Others have pointed out a variety of additional channels through which the cost of adjusting to a new GPT can show up at the macroeconomic level. Greenwood and Yorukoglu (1997) argue that real resources are used up in learning to operate the new GPT. Aghion and Howitt (1998b) point out that the process of reallocating labor from sectors using older technologies to those using the new GPT may involve a rise in unemployment, for the same reason that any large reallocation of labor often entails unemployment in a less than frictionless economic system. Howitt (1998) calibrates to U.S. data a Schumpeterian model with capital-embodied technological change, and shows numerically that the speedup in the rate of innovation induced by a new GPT can reduce

[5] Carlaw, Lipsey and Bekar (2005) develop a comprehensive analysis of economic growth based on general purpose technologies.

the rate of output growth by increasing the rate of induced capital obsolescence, both human and physical. In this calibration, the introduction of a new GPT that raises the productivity of R&D by 50 percent until overall productivity has doubled will reduce the level of per-capita GDP below the path it would otherwise have followed for a period of about two decades, before eventually resulting in a level of GDP twice as high as it would otherwise have been.

A full account of how an economy copes with these adjustments is something that goes beyond incentives, and involves the institutional mechanisms that determine the extent to which the economy is a self-regulating mechanism. This is because the more often an economy is disturbed by major shocks that require people to learn new patterns of behavior, the harder it is for the "invisible hand" to keep it near a harmonious state of smoothly coordinated plans and actions. That is, the more unlikely it is that the classical stability hypothesis implicit in the neoclassical synthesis will be valid.

The self-regulating mechanisms of a modern free-market economy like that of any OECD country are obviously very powerful and robust, because they manage to coordinate the activities of millions of independent transactors, at least most of the time. At the microeconomic level, surpluses and shortages are relatively rare, small and short-lived. At the macro level the system seems to maintain itself within five or ten percent of a full-employment growth path, except for a few dramatic exceptions such as the Great Depression. But surely there are limits to the power of any regulatory mechanism, not matter how skillfully designed or how far evolved, to cope with unusually large and frequent shocks.[6] One of the big challenges that economic growth poses to economic theory is to understand how the regulatory mechanisms of a modern economy work, what their limitations are, and what kinds of collective interventions might be needed to help them cope with circumstances that challenge their efficacy.

All of these questions remain largely unanswered. Almost all of modern economic theory proceeds by assuming that they do not need to be addressed, for it starts from the unexamined premise that observed prices and quantities are generated by a system in equilibrium. In micro theory the convention is to assume that Nash equilibrium prevails in static contexts, or that some form of subgame perfect equilibrium prevails in dynamic settings. In either case, everyone's actions generally depend on expectations of everyone else's and the assumption is that at every node of the tree there are no surprises, in the sense that everyone does what everyone had

[6] Cf. Leijonhufvud (1973).

expected they would do if this state of the world were to prevail. In macro theory the analogous convention is to assume that the economy is always in a rational-expectations equilibrium, where again there are no surprises given the state of the world. It is now widely understood that to assume rational expectations is to assume not just that people are efficient users of information but also that their expectations are perfectly coordinated. My actions in any state of the world will depend on my expectations in that state. For everyone to have anticipated those actions correctly their expectations must have been consistent with mine. How people could acquire a mutually consistent set of expectations is something that we typically don't ask. We just assume they have them.

There have been attempts, in both micro and macro theory to examine the disequilibrium foundations of those equilibrium notions. For example there was a literature on the stability of general equilibrium that flourished in the 1950s and 1960s. But nothing in that literature in any way establishes a presumption of stability. All that can be shown is that there are hypothetical sufficient conditions for stability, such as universal gross substitutability. When theorists discovered what a messy subject they had on their hands they just dropped it, although they had hardly begun to deal with expectations. In fact, most of the literature analyzes only nonmonetary economies in which no one has to trade until the auctioneer has succeeded in arriving at an equilibrium, that is, economies in which effective demand, unemployment, bankruptcy, debt-deflation, endogenous money supply, and so forth have no meaning.

There is also a macroeconomic literature on the stability of full-employment equilibrium, going back to the famously neglected chapter 19 of Keynes's General Theory. Thus Tobin (1947, 1975) and Patinkin (1948) both supported Keynes's view that adverse distributional and expectational effects were likely to make it difficult for an economy to converge upon full employment through the unaided market forces of wage and price adjustment. In recent years it has come to be recognized that the stability of a rational-expectations equilibrium depends on the convergence of a self-referential learning process in which peoples' attempts to learn about a system lead them to take actions that effectively change the system itself. Several years ago (Howitt, 1992) I argued that whether or not this process would converge would depend on the nature of the monetary policies being pursued, and in particular that convergence would require the monetary authority to obey what has subsequently come to be called the Taylor Principle of making the nominal interest rate rise more than point-for-point when inflation increases. This has been shown by a subsequent literature (recently summarized by Evans and Honkapohja, 2001; and Woodford, 2003) to be a valid proposition about the stability of equilibrium under a wide variety of different assumptions.

But all of this work is in its infancy and none of it has reached the position of accepted wisdom, judged by the fact that it has not filtered down to introductory economics textbooks, which are filled with stories of perfectly coordinated individual choices and have nothing to say about how those choices come to be coordinated. Thus it appears that the long-run wealth of a nation depends to a large extent on the convergence properties of a regulatory mechanism about which we as economists know very little.

Moreover, there are good reasons for thinking that policies and institutions that raise the pace of technological progress make it less likely that the mechanism will converge. This is not just because of the increased frequency and amplitude of the shocks with which the system must cope, and not just because of the dangers of financial bubbles and crashes that seem inevitably to be associated with major technological developments,[7] but also because the process of economic growth brings with it a deeper coordination problem that has not yet been addressed in the endogenous growth literature, one which lies at the heart of the growth process.

A particular form of this problem is what motivated Harrod (1939, 1948) and Domar (1946, 1947) to make the contributions that originally gave rise to the modern literature on economic growth. This "Harrod-Domar" problem is the problem of how to ensure enough effective demand so that the increased productive potential created by economic growth will be fully utilized, rather than becoming excess capacity and causing unemployment. It is a question of coordinating the expectations of investors with the yet unarticulated future demands of savers. As long as the marginal propensity to consume is less than unity, business firms will somehow have to see it in their interests to increase their investment outlays each year, and by just the right amount. Harrod rightly perceived that this brought into question the stability of equilibrium. Under his assumptions, any time entrepreneurs found they had overestimated the growth of final sales, they would scale back their collective investment outlays, and the subsequent multiplier effects of this cutback would cause actual sales to fall even more than anticipated. A vicious circle would be created, whereby shortfalls in investment demand would feed on themselves in cumulative fashion.

[7] On this, see Minsky (1992) and Nabar (2004).

One response to this problem is to invoke the classical stability hypothesis – to say that if entrepreneurial expectations don't respond appropriately, then sooner or later wages will have to fall, and the problem will go away. But this response begs the further questions of whether recovery will really be promoted by a debt deflation that will drive many firms out of existence, possibly bringing down with them some of the financial intermediaries whose services will be needed to finance adjustment, whether it will be possible for central banks preoccupied with exchange rates, and controlling a shrinking fraction of the means of payment, to avoid a monetary contraction once prices start falling, and what will counteract the destabilizing expectational and distributional effects upon which Keynes rested his instability case in the General Theory.

As Fazzari (1985) and Sen (1960) have made clear, the Harrod-Domar problem is a particularly intractable one because it involves a positive feedback loop between expectations and outcomes. That is, under the assumptions of the model if entrepreneurs are overly pessimistic in their growth expectations – expecting a rate of growth less than the economy's equilibrium (in Harrod's terms "warranted") rate of growth – then the simple investment multiplier of the Keynesian-Cross model implies they will experience an actual rate of growth even less than they were expecting. In other words, the interactions involved in the multiplier process are such that entrepreneurs will be receiving the wrong signal. Instead of learning that they were overly pessimistic they will learn that they were too optimistic. Any sensible attempt to correct this expectational error will lead them to reduce their expectations by even more, thus leading the economy even further from its equilibrium.

I know of no modern attempt to resolve this Harrod-Domar problem. The literature starting with my 1992 contribution and recently summarized by Woodford would seem to imply that as long as the monetary authority obeys the Taylor Principle the economy should be able to converge to its rational-expectations equilibrium. But my own recent, as yet unpublished, research shows that this is not the case, that when the economy's capacity output is growing then this principle is still necessary but no longer sufficient for stability of equilibrium. Instead the monetary authority must generally also react with sufficient vigour to changes in the level of output, not just to the rate of inflation.

Moreover, the aggregate stability problems that Harrod raised constitute the tip of an iceberg, because adjustment to technological change requires far more than the right level of overall investment demand. We know that Engel curves are not straight lines through the origin. As incomes grow, marginal expenditures are devoted to new and different goods. Full adjustment in a multi-good economy requires entrepreneurs to create the sort of productive capacity and the sort of jobs, in many cases to create entirely new goods and markets, that will enable them ultimately to satisfy the yet unknown wants that people will have when their incomes are higher. Until people have that increased income, or at least enough of a prospect of increased income that they are induced to run down their liquid assets even faster, how are they to make their demands effective, especially if technological change has made them unemployed?

Entrepreneurs not only have to anticipate demands that have not yet been articulated, they have to anticipate the decisions that other entrepreneurs are making, because paying the setup cost of hiring people and capital and developing a market to produce and sell any particular range of goods will only pay off if that range is compatible with the standards, techniques, and strategies that others are developing. And of course these decisions have to be coordinated somehow with those of the unemployed and young workers trying to choose occupations, find sectors, and acquire skills to anticipate the job opportunities of the future.

More generally, in order to accomplish the social objective of exploiting an increased productive potential each year, new trading relationships have to be established that involve literally millions of people. How are these arrangements going to be made when none of the transactors can possibly have a detailed understanding of what is going on, none of them is in direct communication with all the others, and all of them are guided by purely private interests? What signals are going to induce business firms collectively to provide the kind of capital equipment, job opportunities, products, processes and markets that will profitably absorb the potential increases in purchasing power wrought by technological change? How much time, bankruptcy, mismatch and unemployment will it take? Or will adjustment ever be complete without some form of collective guidance, and if so what kind?

5. Looking ahead

I conclude by elaborating on what I said in the introduction, and what should by now be apparent, namely that the coordination issues raised by economic growth are ideally suited for investigation by computational methods. Indeed the computer has already been used by various authors to address some of these questions, mostly by writers in the evolutionary tradition pioneered by Nelson and Winter (1982),[8] but there is much more to be done.

One reason for turning to the computer is that when aggregate outcomes differ from individual intentions it is typically because of a complex set of interactions that are hard to characterize in analytical terms. To illustrate, the above-mentioned result of Aghion, Harris, Howitt and Vickers (2001) to the effect that weaker intellectual property protection would, up to some point, raise aggregate R&D even though it would always have a negative effect on the R&D of a firm in any given situation depended on how the steady-state cross-industry distribution of technology gaps between leading and lagging firms reacted to parameter changes. Except in very special cases the behavior of this distribution was just too complicated for us to sign the comparative-static effect analytically. But the parameter space was simple enough that we were able to demonstrate numerically with reasonable certainty that the effect was always present. And this model was an extremely simple one, with exactly two firms in each industry and all industries ex-ante identical. We really need to examine richer models to test the robustness of such results. The complex web of externalities that growth theory has uncovered makes it highly unlikely that as we go to even richer models we will be able to dispense with the computer for the purpose of discovering robust comparative-static effects.

Another reason for going to the computer in growth theory is to get an idea of the likely size of different effects. Thus in Howitt (1998) I was able to state analytically under what conditions there would be a downturn in overall economic activity following the introduction of a new GPT. But it was only through computational calibration methods that I was able to argue that this is an effect likely to last for many years rather than just a few weeks. These results would have to be replicated in much richer models before they could become generally accepted as true. Again there is no way to do this without computational methods.

The biggest challenge posed by all of these coordination problems is to characterize the mechanisms that keep a modern economic system reasonably near a fully coordinated state most of the time, and hence to deal with the generalized Harrod-Domar problem. We can deal analytically with the stability properties of two-dimensional, sometimes even three-dimensional systems, but beyond this we are lost without the computer.

[8] A good sample of this literature can be found in the book by Dosi *et al.* (1988).

In addition to the issue of dimensionality, no study of the coordination properties of an economic system will be fully satisfactory if it does not come to grips with the elementary fact that most transactions in actual economies are coordinated not by some unspecified agent like the Walrasian auctioneer but by an easily identified set of agents; namely, specialist trading enterprises. Economic transactions do not take place on a do-it-yourself basis but always involve such agents as grocers, department stores, realtors, car dealers, legal firms, accounting firms, and so forth. These are specialist traders that reduce the costs of search, bargaining and exchange, by using their expertise and by setting up trading facilities that enable non-specialists to trade on a regular basis. Collectively they coordinate the exchange process, for better or worse, by setting prices, holding buffer-stock inventories, announcing times of business, entering into implicit or explicit contracts with customers and suppliers, and taking care of logistical problems that arise in delivery, inspection, payment, and other aspects of the transaction process. When there are imbalances between demand and supply, specialist traders typically are responsible for making whatever adjustments are needed to ensure that non-specialists can continue their activities with minimal interruption. Those that do the job poorly do not survive competition.

The job that these trading specialists perform is the "procurement process" that Tesfatsion (2005) argues ACE modeling is ideally designed to study. Howitt and Clower (2000) show how the ACE approach can be used to study the formation and performance of a network of such specialists. In that paper Clower and I imagined a world with a large number of people who could potentially benefit from trading with one another but who lacked the information and the organizational infrastructure necessary to realize those benefits. We asked what would happen if some of the people from time to time were inspired to set up a trading facility, or "shop" that others could use, from which the shopkeeper might also profit by charging different buying and selling prices.

We realized early on that the only sensible approach to modeling how a coordination network might evolve from such a foundation was to write a computer program. For we did not want to impose on people any beliefs or information that implied some kind of prior coordination. Instead we wanted coordination to emerge from the basic assumptions of the model. Thus we needed a model that specified what would happen from any conceivable initial position, no matter what sorts of expectations people started with and no matter how incompatible their plans were to begin with. In short, our model had to constitute a multi-agent system that would generate observed outcomes from any given initial position. This to me is the essential characteristic of ACE methodology that distinguishes it from other uses of the computer in economic theory; in other uses computer programs approximate the behavior of a model, whereas with the ACE approach the program is the model. Since we were going to create a program anyway it seemed sensible to run it on the computer and study its behavior directly rather than seek what would at best be a partial and not very helpful analytical characterization of its properties.

What Clower and I discovered was that even though no one in the world we were describing ever possessed a reliable model of the overall system in which they were participating, nevertheless their interactions often resulted eventually in the emergence of a stable set of shops, each with a stable set of customers and suppliers, and everyone engaging in a pattern of exchange that can be described as a general (Nash) equilibrium in prices and quantities. Moreover, what we found was that whenever such a stable pattern emerged it took on a monetary structure. That is, one of the commodities traded would emerge as a universal medium of exchange, used in every single transaction in the economy, even by people that had no direct use for the commodity and were not capable of producing it.

The fact that this particular application of ACE methodology is capable of growing[9] a coordination network which is sometimes capable of leading people into an equilibrium pattern of exchange, at least in the very simple setting that we postulated, and is also capable of growing some of the ancillary institutions of real-world economic systems, such as monetary exchange, gives me hope that the methodology will some day be capable of shedding light on the big coordination issues raised by economic growth.

[9] Epstein (2005) elaborates on the use of ACE methodology to explain real-world institutions and behavioral patterns by "growing" them on the computer.

References

Acemoglu, D. (2002), "Technical change, inequality and the labor market", Journal of Economic Literature 40:7-72.

Aghion, P., C. Harris, P. Howitt and J. Vickers (2001), "Competition, imitation and growth with step-by-step innovation", Review of Economic Studies 68:467-492.

Aghion, P. and P. Howitt (1994), "Growth and unemployment", Review of Economic Studies 61:477-494.

Aghion, P. and P. Howitt (1998a), Endogenous Growth Theory (MIT Press, Cambridge, MA).

Aghion, P. and P. Howitt (1998b), "On the macroeconomic effects of major technological change", in: E. Helpman, ed., General Purpose Technologies and Economic Growth (MIT Press, Cambridge, MA) pp. 121-144.

Aghion, P., P. Howitt and G.L. Violante (2002), "General purpose technology and wage inequality", Journal of Economic Growth 7:315-345.

Arrow, K.J. (1969), "Classificatory notes on the production and transmission of technological knowledge", American Economic Review Papers and Proceedings 59:29-35.

Arthur, W.B. (1989), "Competing technologies, increasing returns, and lock-in by historical events", Economic Journal 99:116-131.

Carlaw, K.I., R.G. Lipsey and C. Bekar (2005), Economic Transformations: General Purpose Technologies and Sustained Economic Growth (Oxford University Press, Oxford).

Clower, R. and P. Howitt (1998), "Keynes and the classics: An end of century view", in J.C.W. Ahiakpor, ed., Keynes and the Classics Reconsidered (Kluwer, Boston) pp. 163-178.

Coe, D.T. and E. Helpman (1995), "International R&D spillovers", European Economic Review 39:859-887.

Coe, D.T., E. Helpman and A.W. Hoffmaister (1997), "North-south R&D spillovers", Economic Journal 107:134-149.

David, P. (1990), "The dynamo and the computer: An historical perspective on the modern productivity paradox", American Economic Review Papers and Proceedings 80:355-361.

Dawid, H. (2005), "Agent-based models of innovation and technological change", this Handbook.

Domar, E.D. (1946), "Capital expansion, rate of growth and employment", Econometrica 14:137-147.

Domar, E.D. (1947), "Expansion and employment", American Economic Review 37:34-55.

Dosi, G., C. Freeman, R. Nelson, G. Silverberg and L. Soete, eds. (1988), Technical Change and Economic Theory (Pinter Publishers, London).

Easterly, W. (2001), The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics (MIT Press, Cambridge, MA).

Epstein, J.M. (2005), "Remarks on the foundations of agent-based generative social science", this Handbook.

Evans, G. and S. Honkapohja (2001), Learning and Expectations in Macroeconomics (Princeton University Press, Princeton, NJ).

Fazzari, S.M. (1985), "Keynes, Harrod and the rational expectations revolution", Journal of Post Keynesian Economics 8:66-80.

Freeman, C. and C. Perez (1988), "Structural crises of adjustment, business cycles and investment behaviour", in: G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete, eds., Technical Change and Economic Theory (Pinter Publishers, London) pp. 38-66.

Greenwood, J. and M. Yorukoglu (1997), "1974", Carnegie-Rochester Conference Series on Public Policy 46:49-95.

Harrod, R.F. (1939), "An essay in dynamic theory", Economic Journal 49:14-33.
Harrod, R.F. (1948), Towards a Dynamic Economics (Macmillan, London).
Helpman, E. and M. Trajtenberg (1998), "A time to sow and a time to reap: Growth based on general purpose technologies", in: E. Helpman, ed., General Purpose Technologies and Economic Growth (MIT Press, Cambridge, MA) pp. 55-83.
Howitt, P. (1992), "Interest rate control and nonconvergence to rational expectations", Journal of Political Economy 100:776-800.
Howitt, P. (1998), "Measurement, obsolescence, and general purpose technologies", in: E. Helpman, ed., General Purpose Technologies and Economic Growth (MIT Press, Cambridge, MA) pp. 219-251.
Howitt, P. and R. Clower (2000), "The emergence of economic organization", Journal of Economic Behavior and Organization 41:55-84.
Kirman, A. (1992), "Whom or what does the representative individual represent?" Journal of Economic Perspectives 6:117-136.
Klenow, P.J. and A. Rodríguez-Clare (1997), "The neoclassical revival in growth economics: Has it gone too far?" in: B. Bernanke and J. Rotemberg, eds., NBER Macroeconomics Annual (MIT Press, Cambridge, MA) pp. 73-103.
Leijonhufvud, A. (1973), "Effective demand failures", Swedish Journal of Economics 75:27-48. (now Scandinavian Journal of Economics).
Lucas, R.E. Jr. (1972), "Expectations and the neutrality of money", Journal of Economic Theory 4:103-124.
Lucas, R.E. Jr. (2002), Lectures on Economic Growth (Harvard University Press, Cambridge, MA).
Mankiw, N.G., D. Romer and D.N. Weil (1992), "A contribution to the empirics of economic growth", Quarterly Journal of Economics 107:407-437.
Minsky, H.P. (1992), "The financial instability hypothesis", Unpublished (The Jerome Levy Economics Institute, Bard College, New York).
Mokyr, J. (1990), The Lever of Riches (Oxford University Press, Oxford).
Nabar, M. (2004), "Technology-driven booms and crashes, productivity growth and volatility", Unpublished (Brown University, Providence, RI).
Nelson, R. and S. Winter (1982), An Evolutionary Theory of Economic Change (Harvard University Press, Cambridge, MA).
Patinkin, D. (1948), "Price flexibility and full employment", American Economic Review 38:543-564.
Ricardo, D. (1821), The Principles of Political Economy and Taxation, 3rd ed. (Everyman's Library, New York).
Rosenberg, N. (1963), "Technological change in the machine tool industry, 1840-1910", Journal of Economic History 23:414-443.
Rosenberg, N. (1994), Exploring the Black Box: Technology, Economics, and History (Cambridge University Press, Cambridge, UK).
Sargent, T.J. (1993), Bounded Rationality in Macroeconomics (Clarendon Press, Oxford).
Schelling, T.C. (1978), Micromotives and Macrobehavior (Norton, New York).
Sen, A. (1960), "Introduction", in: A. Sen, ed., Growth Economics (Penguin, Harmondsworth, UK) pp. 9-40.
Tesfatsion, L. (2005), "Agent-based computational economics: A constructive approach to economic theory", this Handbook.
Tobin, J. (1947), "Money wage rates and employment", in: S.E. Harris, ed., The New Economics (Albert A. Knopf, New York) pp. 572-587.

Tobin, J. (1975), "Keynesian models of recession and depression", American Economic Review Papers and Proceedings 65:195-202.
Woodford, M. (2003), Interest and Prices: Foundations of a Theory of Monetary Policy (Princeton University Press, Princeton, NJ).
Young, A.A. (1928), "Increasing returns and economic progress", Economic Journal 38:527-542.
Zachariadis, M. (2003), "R&D, innovation, and technological progress: A test of the Schumpeterian framework without scale effects", Canadian Journal of Economics 36:566-586.