11. The renaissance of economic growth research

> No previous episode of enrichment approaches modern economic growth – not China or Egypt in their primes, not the glory of Greece or the grandeur of Rome. (McCloskey, 1994)

> Of all the policy questions concerning growth, the most fundamental is whether there are any policies that an omniscient, omnipotent, benevolent social planner could implement to raise the welfare of all individuals in an economy. (P. Romer, 1989)

11.1 Introduction

In 2002 there were 192 internationally recognized independent countries in the world that included the 191 members of the United Nations plus Taiwan (Alesina and Spolare, 2003). Among these countries are some that by historical standards are extremely rich and many more that are relatively poor. Their size, measured by geographical area, by total population, or by total GDP, varies enormously. Measured by geographical and population size we observe large poor countries (India) and large rich countries (USA) as well as small rich countries (Switzerland) and small poor countries (Sierra Leone). We also see every other possible combination in between. When we examine the economic growth performance of these countries we also see a wide variety of experience, from the high positive rates observed during the last four decades of the twentieth century among the Asian Tigers (Hong Kong, South Korea, Singapore, Taiwan) to the negative growth rates experienced in many sub-Saharan countries during the last couple of decades. Since sustained economic growth is the most important determinant of living standards, there is no more important issue challenging the research efforts of economists than to understand the causes of economic growth. In reviewing the differential growth performances of countries such as India, Egypt, the 'Asian Tigers', Japan and the USA, and the consequences of these differentials for living standards, Lucas (1988) observed that 'the consequences for human welfare involved in questions like these are simply staggering. Once one starts to think about them, it is hard to think about anything else'.

In this chapter we review many of the important theoretical, empirical and political-economy issues relating to modern economic growth and the deter-

minants of living standards that have captivated the research interests of both economists, economic historians and other social scientists during the twentieth century.

11.2 The 'Great Divergence'

Nothing in human history compares with the impact that the Industrial Revolution has had on the living conditions for the world's population. Sustained growth of both total GDP and GDP per capita date from the great transformation unleashed by this event. While economic historians continue to debate the origins, timing and quantitative aspects of the Industrial Revolution, there is no doubt that during the last 250 years the main consequence of this event has been a distinctive regime change as the world economy began to experience a new epoch of what Kuznets (1966) called 'modern economic growth'. It should also be noted that modern economic growth and capitalism are synonymous, for as Baumol (2002) writes:

> what is clear to historians and laypersons alike is that capitalism is unique in the extraordinary growth record is has been able to achieve in its recurring industrial revolutions that have produced an outpouring of material wealth unlike anything previously seen in human history.

Baumol's point had been recognized by Karl Marx and Friedrich Engels over 150 years earlier in their *Communist Manifesto* of 1847 when they observed that 'the bourgeoisie has created more massive and more colossal productive forces than have all preceding generations together'.

In contrast to most of human history, the modern epoch has been characterized by a population explosion, rising life expectancy, rapid urbanization, diversified patterns of employment and steadily rising income per capita for the world as a whole (Easterlin, 1996). However, because the Industrial Revolution and economic growth have spread unevenly across the world, the modern era of human history has also witnessed the emergence of unprecedented global inequality. Since the beginning of the nineteenth century the world economy has experienced what Pomeranz (2000) calls the 'Great Divergence' and Pritchett (1997) refers to as 'Divergence, Big Time'. International differences in living standards, measured by real income per capita, are enormous even after making adjustments to the estimates that take into account variations in purchasing power and household production.

The 'Great Divergence' of income per capita is a modern phenomenon. Before the nineteenth century, for the vast majority of the economies and peoples of the world, the process of economic growth was 'sporadic and inconsistent'. It was not until the second half of the twentieth century that growth spread to many living in the Third World. As Table 11.1 shows, living

Table 11.1 Level[a] and rate of growth[b] of GDP per capita: world and major regions, 0–1998 AD

Region	*0[a]*	*1000[a]*	*1820[a]*	*1998[a]*	*0–1000[b]*	*1000–1820[b]*	*1820–1998[b]*
Western Europe	450	400	1 232	17 921	–0.01	0.14	1.51
Western off-shoots	400	400	1 201	26 146	0.00	0.13	1.75
Japan	400	425	669	20 413	0.01	0.06	1.93
Average, Group A	443	405	1 130	21 470	–0.01	0.13	1.67
Latin America	400	400	665	5 795	0.00	0.06	1.22
Eastern Europe & former USSR	400	400	667	4 354	0.00	0.06	1.06
Asia	450	450	575	2 936	0.00	0.03	0.92
Africa	425	416	418	1 368	–0.00	0.00	0.67
Average, Group B	444	440	573	3 102	–0.00	0.03	0.95
World	444	435	667	5 709	–0.00	0.05	1.21

Notes:
[a] Measured in 1990 international dollars.
[b] Annual average compound growth.

Source: Maddison (2001), Table 1.2.

standards across the world, as measured by GDP per capita, did not improve in any significant way during the first one thousand years AD. However, since about 1750, beginning in Great Britain, the phenomenon of sustained modern economic growth has become 'the defining feature of human history' and by 1950 embraced a third of the population of the earth.

Contemporary differences in living standards are themselves the product of differences in growth rates that have been observed during the last 200 years and are highlighted in Table 11.2. Three important commonly used measures of living standards are represented, namely gross national income (GNI) per capita measured in international dollars (PPP – purchasing power parity – $), life expectancy, and the Human Development Index (HDI). The HDI is a composite measure of three equally weighted basic components, namely, real income per capita (PPP US$ = *Ypc*) adjusted to reflect the assumption of rapidly diminishing marginal utility of income above the world average; longevity as measured by life expectancy at birth (*L*); and educational attainment (*E*) captured by the adult literacy rate (weighted 2/3) and the combined gross primary, secondary and tertiary enrolment ratio (weighted 1/3). Therefore, the HDI estimate for any economy (*j*) is a simple weighted average of *Ypcj* + *Lj* + *Ej*. Although there are serious index number problems in using the HDI as a measure of living standards and the HDI has come in for a considerable amount of criticism, it has nevertheless proved to be a useful additional development indicator, complementing, but not replacing, the traditional 'commodity'-based measures of progress such as income per capita. While there has been unprecedented divergence between the income per capita of the OECD economies and many developing countries, Crafts (1999, 2000) has argued that a more optimistic picture of the progress of human welfare emerges if we examine long-run trends in the HDI. For example, the HDI scores for many poor countries in 2002 are well ahead of the estimated 1870 HDI scores for the leading countries of that time (current G7 countries) as measured by their per capita income. Crafts concludes that by taking a broader view of progress, it is 'likely that the growth of living standards since 1870 as measured by real national income per capita is substantially underestimated' (Crafts, 2001; see also Becker et al., 2003; Crafts, 2003).

Data on total population for each country is also included in Table 11.2. The 40 countries included accounted for 4795.7 million (79 per cent) of the world's population of 6054 million in 2000. Important points to note include the close correlation between income per capita, life expectancy and the HDI; the underperformance of Botswana and South Africa on their life expectancy and HDI scores relative to their position in the income per capita ranking; the ratio of the USA's GNI per capita to Sierra Leone's is a staggering 72–1 (note, the highest GNI per capita recorded by the World Bank for 2002 was Luxembourg, with 51,160 PPP$).

Table 11.2 Three indicators of living standards: 40 countries

Country	*GNI per capita (PPP$), 2002*[a]	*Life expectancy in years, 2000*[b]	*HDI, 2000*[b]	*Total population (millions), 2000*[c]
USA	35 060	77.0	0.939	282
Canada	28 070	78.8	0.940	31
Australia	26 960	78.9	0.939	19
Hong Kong	26 810	79.5	0.888	4.4[d]
Germany	26 220	77.7	0.925	82
France	26 180	78.6	0.928	59
Japan	26 070	81.0	0.933	127
UK	25 870	77.7	0.928	60
Italy	25 320	78.5	0.913	58
Singapore	23 090	77.6	0.885	2.3[d]
Korea, Rep.	16 480	74.9	0.882	47
Czech Rep.	14 500	74.9	0.849	10
Hungary	12 810	71.3	0.835	10
Saudi Arabia	11 480	71.6	0.759	21
Poland	10 130	73.3	0.833	39
Argentina	9 930	73.4	0.844	37
South Africa	9 870	52.1	0.695	43
Chile	9 180	75.3	0.831	15
Mexico	8 540	72.6	0.796	98
Malaysia	8 280	72.5	0.782	23
Russian Fed.	7 820	66.1	0.781	146
Botswana	7 770	40.3	0.572	2
Brazil	7 250	67.7	0.757	170
Thailand	6 680	70.2	0.762	61
Iran	6 340	68.9	0.721	64
Turkey	6 210	69.8	0.742	65
Ukraine	4 650	68.1	0.748	50
China	4 390	70.5	0.726	1 261
Egypt	3 710	67.3	0.642	64
Indonesia	2 990	66.2	0.684	210
India	2 570	63.3	0.577	1 016
Vietnam	2 240	68.2	0.688	79
Zimbabwe	2 120	49.2	0.551	12
Pakistan	1 940	60.0	0.499	138
Bangladesh	1 720	59.4	0.478	130
Kenya	990	50.8	0.513	30
Nigeria	780	51.7	0.462	127
Ethiopia	720	43.9	0.327	64
Tanzania	550	51.1	0.440	34
Sierra Leone	490	38.9	0.275	5
World[e]	7 570 (A)	66.9 (A)	0.722 (A)	4 795.7 (T)

Notes:

[a] Gross national income per capita, PPP$, *World Development Indicators, 2003*, World Bank.

[b] Human Development Index and Life Expectancy, *Human Development Report 2002*, United Nations.

[c] Total population, *World Development Report 2002*, World Bank.

[d] Hong Kong and Singapore data from *Human Development Report 2002*.

[e] A = Average, T = Total.

11.3 In Praise of Economic History

It is our view that a knowledge of history in general, and economic history in particular, is important to understand how societies and economies change. Since contemporary economic historians are primarily interested in the long-run development of economies, they seek to understand the fundamental causes of economic growth, the determinants of technological progress, the evolution and impact of institutions, and the historical origins of current economic problems. With respect to the determination of technological progress, which is now at the forefront of endogenous growth theory, Wright (1997) argues that if economists wish to take technology seriously then economics 'will have to become a more historical discipline' (see also Mokyr, 2002, 2005). In his influential 1986 paper, William Baumol also advises economists interested in long-run growth to pay more attention to the 'brilliant insights' and 'powerful analysis' of economic historians.

Before the 1960s, even though large amounts of invaluable quantitative data were produced by economic historians, much of their analysis tended to be atheoretical. The traditional approach to the study of economic history was largely descriptive. This situation changed dramatically during the 1960s as formalization and analytical rigour spread from mainstream economics to the field of economic history. From the early 1960s scholars such as Nobel Memorial Laureates Robert Fogel and Douglass North pioneered the 'new' quantitative approach to economic history, or 'cliometrics', defined as 'the application of economic theory and quantitative methods to the study of history' (Goldin, 1995).

During the last four decades the 'cliometric revolution' has demonstrated that economic historians have much to gain from a knowledge of economic theory and methodology. The 'new' approach has emphasized the need for scholars to be precise and explicit about which hypotheses are being tested in order to connect the historical investigation with quantitative analysis. However, economists also have much to gain from a greater knowledge of history, in particular economic history (Snowdon, 2002c). Indeed, one of the main developments highlighted in this chapter is how in recent years economic theorists appear to have followed Goldin's (1995) advice that 'only the oblivious can ignore history in modern economics, and only the unenlightened would chose to do so'. Goldin's recommendation is especially relevant to those economists interested in the long-run issue of economic growth. Not only does the past provide a gigantic laboratory for testing various hypotheses in economics; history also contains many lessons that can provide useful information for contemporary policy makers, not least those in the developing countries and transition economies. Because the past shapes the present it must also influence the future. As Goldin argues, the 'remnants of the past,

which shape the realm of the possible today, are always with us, norms, structures, institutions, and even people'. While history rarely repeats itself exactly, it does offer guidance, broadens our stock of knowledge, highlights what may be important in determining outcomes, and 'enables us to identify and read signals' (Horrell, 2003). It is worth noting that several prominent macroeconomists have, in recent years, made important contributions to growth analysis by engaging in quantitative economic history (see Lucas, 2000b, 2002; Hansen and Prescott, 2002; Parente and Prescott, 2005; and section 11.21).

11.4 Back to the Long Run

Following the stimulating and important contributions of Abramovitz (1986), Baumol (1986), P. Romer (1986) and Lucas (1988), the study of long-run economic growth once again became a very active research area. Increasingly, many economists accept that macroeconomic analysis and policies resting solely on the short view will be, to quote Viner, 'a structure built on shifting sands' (Baumol, 1986). This reorientation of research emphasis is regarded as long overdue by many economists. Although the revival of research into the elusive ingredients of growth began in the mid-1980s it was not until the 1990s that the literature really exploded. Two important surveys of macroeconomics bear witness to this fact. Mankiw's (1990) survey of macroeconomics makes no mention of economic growth whatsoever, while Fischer's (1988) survey devotes a mere four sentences to the 'Theory of Growth' and only refers to Paul Romer's 1986 paper in passing. However, by 1989 Robert Barro and Paul Romer were launching a major joint research project on economic growth through Harvard University and the National Bureau of Economic Research and, reflecting this development, in 1996 the first issue of the *Journal of Economic Growth* was published. After 1990 a constant flow of new books and journal survey articles also began to appear, indicating the much higher profile and new urgency given by economists to research on economic growth (see, for example, Grossman and Helpman, 1991; Fagerberg, 1994; P. Romer, 1994a; Mankiw, 1995; Crafts and Toniolo, 1996; Aghion and Howitt, 1998; Gylfason, 1999; Temple, 1999; Solow, 2000, 2001; Easterly, 2001a; Easterly and Levine, 2001; Jones, 2001a; Lucas, 2002; Barro and Sala-i-Martin, 2003; Rodrik, 2003; Aghion and Durlauf, 2005; see also Snowdon and Vane, 1997a; Snowdon, 2003a).

Given the significant adverse impact that poor growth performance has on economic welfare and the resultant importance attached to growth by economists, it is perhaps surprising that during the twentieth century the research effort in this field has been cyclical. Concern about the sustainability of economic growth was a major concern of the classical economists, with the

pessimism of Thomas Malthus and David Ricardo contrasting with the optimism of Adam Smith (Rostow, 1990). However, during the period 1870–1929 economists' research was heavily influenced by the 'marginalist revolution' and was therefore predominantly micro oriented, being directed towards issues relating to the efficient allocation of given resources. For a quarter of a century after 1929–33, issues relating to the Great Depression and Keynes's response to that event dominated discussion as the new science of macroeconomics evolved. In the period 1939–56 growth theory was dominated by the neo-Keynesian contributions of Roy Harrod (1939, 1948) and Evsey Domar (1946, 1947, 1948), and in the period 1956–70 by the seminal contributions of Nobel Memorial Laureate Robert Solow (1956, 1957), who, along with Trevor Swan (1956), pioneered work on the neoclassical growth model. However, as Domar commented in 1957, 'in economic theory growth has occupied an odd place: always seen around but seldom invited in. It has either been taken for granted or treated as an afterthought'.

It is certainly the case that research on the theoretical front in this field ran into diminishing returns and 'effectively died' in the 1970–85 period. This was due mainly to its perceived lack of empirical relevance, the diversion of economists' research interests towards business cycle analysis in the wake of the aggregate instability experienced throughout the world in the 1970s, and the impact of the rational expectations 'revolution' within academia (see Chapter 5, and Barro and Sala-i-Martin, 2003). In a survey of the contents of three leading economics journals Laband and Wells (1998) found that

> during the first half of the twentieth century, there was increasing scholarly output of papers written about economic growth and development. This scholarly interest/production peaked in the 1950s and then declined markedly throughout the 1960s, 1970s, and 1980s. This trend reversed itself abruptly during the first half of the 1990s, with a surge in production of articles on economic growth.

How far this burgeoning literature has significantly progressed economists' knowledge of the causes of growth remains to be seen (see Nelson, 1997; Abramovitz, 1999; Kenny and Williams, 2001).

During the last 50 years there has been a shift of theoretical focus as neo-Keynesian growth models were replaced by neoclassical models as the dominant framework for analysis. Neoclassical theories have in turn been challenged by endogenous growth theory since the mid-1980s (Snowdon and Vane, 1997a; Solow, 2001). But what explains the high-profile resurgence of interest in growth analysis during the last couple of decades? We identify this resurgence of interest since the mid-1980s to have been stimulated by the following 12 factors, of which the first three played a crucial role:

1. New theoretical insights inspired by the research of Paul Romer (1986, 1987b, 1989, 1990, 1994a) and Robert Lucas (1988, 1990b, 1993, 2002); new theoretical tools have been of paramount importance.
2. The availability of a rich array of new data for a large number of countries (for example Summers and Heston, 1991; Maddison, 2001). Economists now have data for most countries which extend back to 1960. Recent empirical research has also focused on patterns of cross-country growth (Durlauf and Quah, 1999).
3. A growing realization that a large number of developing countries, particularly in sub-Saharan Africa, were not 'catching up' and converging with the levels of income per capita of the rich OECD economies (Abramovitz, 1986, Baumol, 1986, P. Romer, 1986, 1989; Lucas, 1988).
4. The sudden and unexpected collapse of the Soviet Union and other 'Eastern Bloc' economies at the end of the 1980s focused attention on the relationship between social, political and economic structures and an economy's capacity to sustain economic growth (Fukuyama, 1992; Snowdon, 2003b).
5. Increasing concern during the 1980s that the economic position of the USA relative to other major OECD economies, especially Japan and Germany, was being eroded (Thurow, 1992).
6. Concern relating to the causes of the productivity growth slowdown, beginning in the late 1960s/early 1970s, but not clearly recognized until the early 1980s (Fischer et al., 1988; Baumol et al., 1989). Writing in 1988, Fischer described this event as 'the most significant macroeconomic development of the last two decades'. More recently, in the late 1990s interest in the USA has focused on a productivity acceleration associated with the emergence of an information-technology-driven 'new economy' (see Gordon, 2000b; Jorgenson, 2001; Jorgenson and Stiroh, 2000).
7. Increasing awareness of problems relating to the measurement of economic growth and that the true rate of progress is likely to be 'substantially underestimated' using conventional estimation techniques (Fogel, 1999; Nordhaus, 2001). The findings of the Boskin Commission suggest that US GDP growth has been underestimated by about 0.9 per cent per annum in the period 1970–96 (Boskin, 1996). In addition, the rise of information technology and with it the 'knowledge' (or 'weightless') economy, and a potential reversion to non-market production further increased the need to refine national income accounting techniques (Stafford, 1999).
8. Increasing recognition of the spectacular growth performance displayed by the 'East Asian Tiger' economies as well as the 'growth disasters' and disappointments experienced in many developing economies, espe-

cially in sub-Saharan Africa, Latin America and Southern Asia (World Bank, 1993; Bhagwati, 1993; Bloom and Sachs, 1998; A. Taylor, 1998; Collier and Gunning, 1999a, 1999b).

9. The increasing influence, during the 1980s, of the real business cycle approach to the study of economic fluctuations where the Solow neoclassical growth model is used as the benchmark for studying both fluctuations and growth (Kydland and Prescott, 1982). Real business cycle theorists argue that the growth process has a large random element and it is this that causes the 'short-run' fluctuations of output. Aggregate instability is simply the 'manifestation of the process of stochastic growth' (Ryan and Mullineux, 1997). Whereas mainstream macroeconomists treat business cycles and economic growth as largely independent phenomena, real business cycle theorists view aggregate fluctuations as the economy's optimal response to supply-side (productivity) shocks (see Chapter 6).
10. A much-neglected factor in explaining the growth of interest in a particular branch of an academic discipline is the influence exerted by 'internal scientific characteristics'. Because the incentive structure in academia is related closely to publications, particularly in the USA, a new idea or research programme which is 'article-laden', with a rich vein of topics to mine, is highly contagious. An article-laden new theory supported by new data sets, which challenges the existing orthodoxy, will always prove to be a powerful force within academia (Snowdon and Vane, 1996).
11. During the last decade there has been an increasing number of papers devoted to what we would label the 'new political macroeconomics of growth'. In addition to the work of Mancur Olson (1993, 2000), the innovative contributions of Daron Acemoglu and his co-authors have been particularly important in reawakening interest in the 'political barriers' to growth (see Acemoglu and Robinson, 2000a, 2000b, 2000c, 2001, 2003; Snowdon, 2004c).
12. For some economists, such as Robert Lucas (1987, 2003) and Edward Prescott (1996), the renewed interest in growth stems from their belief that business cycle fluctuations 'are not costly to society' and that it is more important for economists to worry about 'increasing the rate of increase in economy-wide productivity and not smoothing business fluctuations' (Prescott, 1996). Indeed, attempts to stabilize the economy in the short term could adversely affect long-term growth prospects (Cooley and Ohanian, 1997; Blackburn, 1999).

11.5 Why is Economic Growth So Important?

Since the middle of the eighteenth century human history has been dominated by the phenomenon of modern economic growth. In the eighteenth and nineteenth centuries economic growth had been largely confined to a small number of countries (Bairoch, 1993; Easterlin, 1996; Maddison, 2001). Gradually, 'modern' economic growth spread from its origins in Great Britain to Western Europe and initially to overseas areas settled by European migrants (Landes, 1969, 1998). The dramatic improvement in living standards that has taken place in the advanced industrial economies since the Industrial Revolution is now spreading to other parts of the world. However, this diffusion has been highly uneven and in some cases negligible. The result of this long period of uneven growth is a pattern of income per capita differentials between the richest and poorest countries of the world that almost defies comprehension (see Tables 11.1 and 11.2). The importance of economic growth as a basis for improvements in human welfare cannot be overstated and is confirmed by numerous empirical studies (see, for example, Dollar and Kraay, 2002a, 2002b). Even small inter-country differences in growth rates of per capita income, if sustained over long periods of time, lead to significant differences in relative living standards between nations. There is no better demonstration of this fact than the impact on living standards of the growth experiences of the 'miracle' East Asian economies compared with those of the majority of sub-Saharan African economies since 1960 by which time the decolonization process was well under way.

It is worth remembering throughout this discussion that the doubling time for any variable growing exponentially at an annual rate of 1 per cent is approximately 70 years. The so-called 'rule of seventy' says that if any variable grows at g per cent per annum, then it will take approximately $70/g$ years for that variable to double in value. More formally, this can be demonstrated as follows (Jones, 2001a). If y_t is per capita income at time t, and y_0 some initial value of per capita income, then the value of y_t is given by equation (11.1):

$$y_t = y_0 e^{gt} \tag{11.1}$$

Equation (11.1) says that if y_0 grows continuously and exponentially at a rate g, its value at time t will be y_t. Let the length of time that it will take for per capita income to double (that is, for $y_t = 2y_0$) be t^*. Therefore t^* will be the solution to equations (11.2) and (11.3) below:

$$2y_0 = y_0 e^{gt*} \tag{11.2}$$

$$t^* = \log 2/g \tag{11.3}$$

Since log 2 ≈ 0.7, then for a growth rate of 1 per cent, $t^* \approx 0.7/0.01 \approx 70$ years. We can generalize this relationship and say, for example, that any country which has per capita income growth of g = 5 per cent will see its living standards double in $70/g = 14$ years. Thus the impact of even small differentials in growth rates, when compounded over time, are striking. David Romer (1996) has expressed this point succinctly as follows: 'the welfare implications of long-run growth swamp any possible effects of the short-run fluctuations that macroeconomics traditionally focuses on'. Barro and Sala-i-Martin (2003) also argue that 'economic growth … is the part of macroeconomics that really matters', a view in large part endorsed by Mankiw (1995), who writes that 'long-run growth is as important – perhaps more important – than short-run fluctuations'.

Table 11.3 illustrates the compounding effect of sustained growth on the *absolute* living standards of five hypothetical countries, labelled A–E, each of which starts out with an income per capita of $1000.

Table 11.3 The cumulative impact of differential growth rates

Period in years	*A* g = *1%*	*B* g = *2%*	*C* g = *3%*	*D* g = *4%*	*E* g = *5%*
0	$1 000	$1 000	$1 000	$1 000	$1 000
10	1 100	1 220	1 340	1 480	1 630
20	1 220	1 490	1 800	2 190	2 650
30	1 350	1 810	2 430	3 240	4 320
40	1 490	2 210	3 260	4 800	7 040
50	1 640	2 690	4 380	7 110	11 470

The data show how, over a period of 50 years, variations in growth rates (g%) between countries A–E, cause a substantial divergence of relative living standards.

The hypothetical data in Table 11.3 are replicated in Figure 11.1, which clearly highlights how diverging living standards can emerge over what is a relatively short historical time period of 50 years. Following Galor and Mountford (2003) in Figure 11.2 we also reproduce the actual growth experience of different regions of the world using Maddison's (2001) data.

Figures 11.1 and 11.2 illustrate how economic growth is the single most powerful mechanism for generating long-term increases in income per capita as well as divergence in living standards if growth rates differ across the

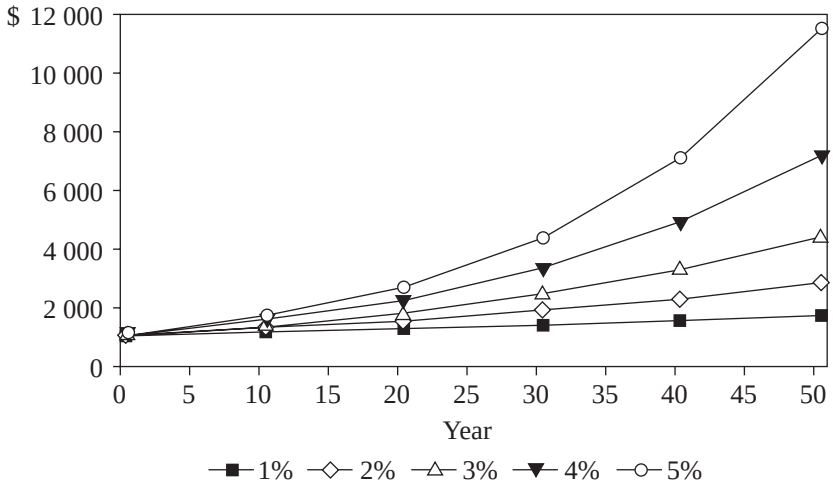
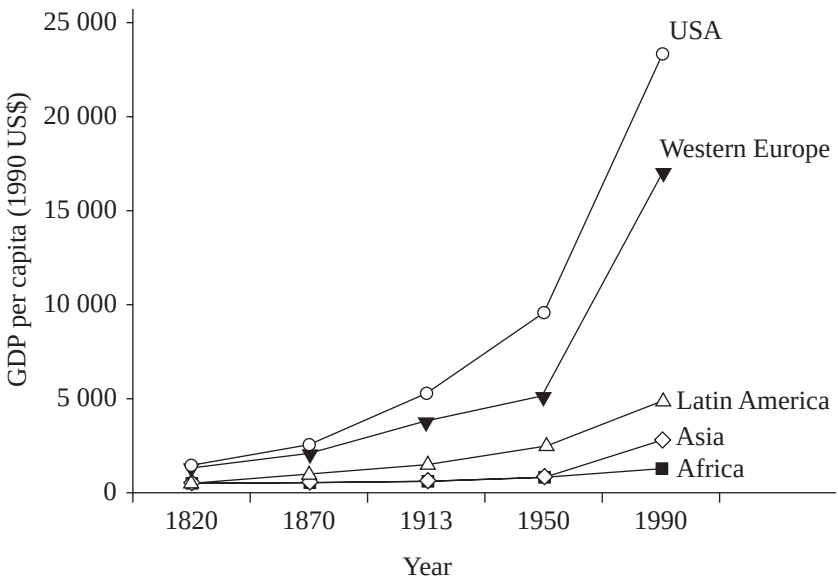


Figure 11.1 The impact on per capita income of differential growth rates



Source: Galor and Mountford (2003).

Figure 11.2 The great divergence

regions and countries of the world. Over very short time horizons the gains from moderate economic growth are often imperceptible to the beneficiaries. However, over generations the gains are unmistakable. As Rosenberg and Birdzell (1986) argue:

> Over the year, or even over the decade, the economic gains (of the late eighteenth and nineteenth centuries), after allowing for the rise in population, were so little noticeable that it was widely believed that the gains were experienced only by the rich, and not by the poor. Only as the West's compounded growth continued through the twentieth century did its breadth become clear. It became obvious that Western working classes were prospering and growing as a proportion of the whole population. Not that poverty disappeared. The West's achievement was not the abolition of poverty but the reduction of its incidence from 90 per cent of the population to 30 per cent, 20 per cent, or less, depending on the country and one's definition of poverty.

These views, emphasizing the importance of long-run economic growth, have also featured prominently in policy statements in recent years. For example, in the concluding section of the introduction to the Council of Economic Advisers' (CEA) *Economic Report of the President* (2004, p. 27) we find the following statement:

> As the Founding Fathers signed the Declaration of Independence the great economist Adam Smith wrote: 'Little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism but peace, easy taxes, and tolerable administration of justice: all the rest being brought about by the natural course of things'. The economic analysis presented in this *Report* builds on the ideas of Adam Smith and his intellectual descendants by discussing the role of government in creating an environment that promotes and sustains economic growth.

The importance of sustainable growth was also emphasized in providing justification for the change in monetary policy arrangements made in the UK in May 1997. Within days of winning the election, the 'New Labour' government announced that the Bank of England was to operate within a new institutional framework giving it operational independence with respect to the setting of short-term interest rates. The economic case for this new arrangement was provided on 6 May 1997 by the Chancellor of the Exchequer, Gordon Brown, when, in an open letter to the Governor of the Bank of England, he stated that 'price stability is a precondition for high and stable levels of growth and employment, which in turn will help to create the conditions for price stability on a sustainable basis. To that end, the monetary policy objective of the Bank of England will be to deliver price stability (as defined by the Government's inflation target) and, without prejudice to this objective, to support the Government's economic policy, including its objectives for growth and employment' (Brown, 1997; see also Brown, 2001).

The power of economic growth to raise living standards is perhaps best illustrated by the history of the twentieth century. Despite two devastating world wars, the Great Depression and collapse of international integration during the interwar period, and the rise and fall of the socialist experiment, the majority of the world's population are better off than their parents and grandparents in terms of income per capita (PPP$). If life expectancy is taken into account there has been a remarkable improvement in welfare (see Crafts, 2003).

11.6 Modern Economic Growth in Historical Perspective

In addition to the 'Great Depression' of the 1930s, and the 'Great Inflation' of the 1970s, the third phenomenon that has dominated the macroeconomic history of the twentieth century has been the spread of economic growth among the economies of the world. Indeed, Robert Lucas (see interview at the end of Chapter 5) believes 'that economic growth, and particularly the diffusion of economic growth to what we used to call the Third World, is *the* major macroeconomic event of the twentieth century'. As Maddison's (2001) data show, before the modern era living standards for the vast majority of the world's population progressed at a glacial pace. In reflecting on the 'Economic Possibilities of Our Grandchildren', Keynes (1930b) commented that 'From the earliest times of which we have record … there was no very great change in the standard of life of the average man living in the civilised centres of the earth … This slow rate of progress, or lack of progress, was due to two reasons – to the remarkable absence of important technical improvements and to the failure of capital to accumulate.'

Since an increase in the capacity to produce can either be absorbed by an increase in population or lead to an increase in per capita income, it is important at the outset to distinguish between extensive and intensive growth. Reynolds (1985) defines extensive growth as a situation where an increase in GDP is fully absorbed by population increase with no upward trend in per capita income. The pre-modern world economy was not characterized by persistent stagnation. The fact that for thousands of years the world's population increased, even if 'glacially slowly', is evidence of extensive growth. If we assume that for the vast majority of people, subsistence living was the norm, then a larger population is only possible if total output also rises (Kremer, 1993). So extensive growth has been 'fairly common' throughout human history (Lal, 1999).

In contrast, intensive growth is where GDP growth exceeds population growth, allowing a sustained rise in living standards as measured by real income per capita. As Reynolds shows, periods of intensive growth have usually been preceded by a long period of extensive growth, often lasting

several centuries, and the significant 'turning point' for any economy is the period of transition from extensive to intensive growth. The 'turning point is actually a period of a decade or two around the cited year, during which one observes a significant and continuing rise in *per capita* income' (Reynolds, 1994). In the past, in predominantly agrarian (organic) economies, the possibilities for sustained intensive growth were extremely limited. The availability and productivity of land determined the amount of extensive growth, but once the supply of suitable agricultural land was exhausted, diminishing returns set in. When these forces are combined with Malthusian population dynamics it is hardly surprising to find that many classical economists predicted the inevitability of a long-run stationary state involving subsistence standards of living for the vast majority of humanity.

It is also useful to make a further distinction when discussing intensive growth. Eric Jones (1988) distinguishes between two forms of intensive growth, namely, 'Smithian growth' and 'Promethian growth'. Smithian intensive growth relies on the gains to productivity that can be made from the division of labour, specialization and trade. Such growth must eventually run into diminishing returns as there are limits to the gains from resource reallocation. In contrast, Promethian intensive growth is sustainable, being driven by technological progress and innovation. It was in the latter part of the eighteenth century that we began to see the emergence of Promethian growth in Britain as during the Industrial Revolution a predominantly organic economy was replaced by a mineral-based one. Of course, the billion-dollar question that many economists and economic historians have tried to answer is: why did Promethian growth begin in a specific geographical location (that is, Britain) and why at a specific time in history? (See Landes, 1969, 1990, 1998; Crafts, 1983, 1985; E. Jones, 1988; Wrigley, 1988; Mokyr, 1990, 1993, 2005; Diamond, 1997; Lal, 1999; Jay, 2000; Pomeranz, 2000, Jones, 2001b.)

The phenomenon of intensive Promethian growth represents, in Easterlin's (1996) view, a distinctive 'regime change'. Easterlin divides world economic history into three epochs, each of which possess distinctive characteristics in terms of the main form of occupation, principal type of population settlement, and the growth rates of population and real GDP per capita. Easterlin's epochs consist of first, a prehistoric epoch ending about 8000 BC; second, an epoch of settled agriculture, initiated by the Neolithic agricultural revolution, which lasted until the middle of the eighteenth century; and third, an epoch of modern economic growth involving an enormous transformation in the structure and character of economic activity. In the modern growth regime, initially, the positive Malthusian relationship between income per capita and population growth persists, leading to a population explosion. Eventually, however, the modern growth regime 'is characterised by steady growth in both income per capita and the level of technology' and this leads to 'a negative relation-

ship between the level of output and the growth rate of population' as the demographic transition kicks in (Galor and Weil, 2000). Several economists have recently argued that any story of the growth process, in addition to accounting for the modern experience of sustained growth, should also be able to account for the long period of Malthusian stagnation (see Galor and Weil, 1999, 2000; Galor and Moav, 2002; Hansen and Prescott, 2002; Parente and Prescott, 2005; see also section 11.21).

11.7 The Stylized Facts of Growth

A convincing theory of economic growth obviously needs to be consistent with the stylized facts of growth that have emerged from historical experience. It was Kaldor (1961) who first set out what he considered to be the main empirical observations with which any growth theory needed to be consistent. Kaldor's six stylized facts, or broad tendencies, are set out below (K1–K6), together with the additional 'facts' (R7–R11) noted by P. Romer (1989) and (J12–J14) noted by Jones (2001a).

K1 Output per worker grows continuously, with no secular tendency for the rate of growth of productivity to decline.
K2 The capital–labour ratio shows continuous growth.
K3 The rate of return on capital is stable.
K4 The capital–output ratio is stable.
K5 The shares of labour and capital in GDP remain stable.
K6 We observe significant variation in the rate of growth of productivity across countries.
R7 In a broad cross-section of countries the average growth rate is uncorrelated with the level of per capita income.
R8 Growth is positively correlated with the volume of international trade.
R.9 Growth rates are negatively correlated with population growth.
R10 Growth accounting research always finds a 'residual'; that is, accumulation of factor inputs alone cannot account for growth.
R11 High-income countries attract both skilled and unskilled workers.
J12 There is enormous variation in income per capita across countries.
J13 Growth rates for the world as a whole, and for individual countries, vary substantially over time.
J14 The relative position of any country in the world distribution of income can change.

Of course not all of these stylized facts are independent. As Romer (1989) points out, fact K2 results from facts K1 and K4. Facts K4 and K5 imply fact K3. Romer also questions the validity of K5 (see also Jones, 2004). With

respect to fact J13, economists have only recently attempted to provide a comprehensive theory that can explain the evolution of growth rates from Malthusian stagnation to 'modern economic growth' (see Galor and Weil, 1999, 2000; Hansen and Prescott, 2002).

11.8 Proximate v. Fundamental Sources of Growth

In Temple's (1999) survey of growth empirics he highlights the fact that one of the important history lessons since 1960 has been that while some countries have succeeded in 'making a miracle', other countries have been 'growth disasters'. When analysing the experience of the miracle economies, economists need to 'use these events to help in assessing economic policies that may affect growth rates in other countries' (Lucas, 1993). However, this cannot be done without having theoretical structures in place that help researchers make sense of the 'mass' of data that is now available to economists. As Lucas argues, to be able to glean valuable lessons from the East Asian experience, 'One needs, in short, a theory'. Echoing this theme, Temple (1999) reminds economists of what is perhaps the key issue: 'Why have some countries grown rich while others remain poor? It is hard to think of a more fundamental question for economists to answer.'

In analysing developments in growth theory it is useful to begin by distinguishing between proximate and fundamental causes of growth. The proximate causes relate to the accumulation of factor inputs such as capital and labour, and also to variables which influence the productivity of these inputs, such as scale economies and technological change. The research of growth accountants such as Denison (1967, 1974, 1985), Jorgensen (1996, 2001) and Maddison (1972, 1987, 1995) has produced a useful taxonomy of the various proximate sources of growth, and the neo-Keynesian, neoclassical and endogenous growth theories tend to concentrate on modelling these proximate variables. However, once we have considered the impact of these proximate determinants of growth we are left with the deeper question: 'Why are some countries so much better than others at accumulating human and physical capital and producing or adopting new ideas and knowledge?' That is, we need to investigate the fundamental determinants of growth (see Rodrik, 2003).

The fundamental or deep sources of growth relate to those variables that have an important influence on a country's ability and capacity to accumulate factors of production and invest in the production of knowledge. For example, Temple (1999) considers the following 'wider' influences on growth: population growth, the influence of the financial sector, the general macroeconomic environment, trade regimes, the size of government, income distribution and the political and social environment. To this list Gallup et al.

(1998) would add the neglected influence of geography. Moving from the proximate to the fundamental causes of growth also shifts the focus of attention to the institutional framework of an economy, to its 'social capability' (Abramovitz, 1986), 'social infrastructure' (Hall and Jones, 1997, 1999) or 'ancillary variables' (Baumol et al., 1994). There is now widespread acceptance of the idea that 'good' governance and institutions and incentive structures are an important precondition for successful growth and development (World Bank, 1997, 2002).

In his historical survey of economic growth analysis, Rostow (1990) put forward a central proposition that 'from the eighteenth century to the present, growth theories have been based on one formulation or another of a universal equation or production function'. As formulated by Adelman (1958), this can be expressed as equation (11.4):

$$Y_t = f(K_t, N_t, L_t, A_t, S_t) \tag{11.4}$$

where K_t, N_t and L_t represent the services flowing from the capital stock, natural resources (geography) and labour resources respectively, A_t denotes an economy's stock of applied knowledge, and S_t represents what Adelman calls the 'sociocultural milieu', and Abramovitz (1986) more recently has called 'social capability', within which the economy functions. More sophisticated models distinguish between human and physical capital. Indeed, many authors regard human capital as the key ingredient of economic growth (Lucas, 1988; Galor and Moav, 2003). Heckman (2003), for example, argues that China's below average spending on investment in education compared to physical capital accumulation is 'a serious distortion' of policy that is likely to retard progress in China. Goldin (2001) has also attributed much of the US economic success in the twentieth century to the accumulation of human capital.

According to Rostow (1990), 'something like the basic equation is embedded equally in Hume's economic essays, Adam Smith's *The Wealth of Nations*, the latest neoclassical growth model, and virtually every formulation in between'. This universal equation encompasses both proximate and fundamental causes of economic growth and Abramovitz drew attention to the importance of these factors 50 years ago (see Nelson, 1997). Clearly, S_t contains the influence of non-economic as well as economic variables which can influence the growth potential and performance of an economy including the institutions, incentives, rules and regulations that determine the allocation of entrepreneurial talent (Baumol, 1990). Hence in recent years economists' research into the 'deeper' determinants of growth has led some to stress the importance of institutions and incentive structures (North, 1990; Olson, 2000), trade and openness (Krueger, 1997; Dollar and Kraay, 2003) and the much-

neglected impact of geography (Bloom and Sachs, 1998). It is important to note that Adam Smith had highlighted all three of these 'deeper' determinants of growth over 200 years ago!

In sections 11.17–11.20 we will examine the 'deeper' determinants of economic growth in more detail, but first, in sections 11.8–11.10, we survey the three main waves of growth theory that have been influential in the second half of the twentieth century to date. All three approaches emphasize the proximate determinants of growth, namely:

1. the neo-Keynesian Harrod–Domar model;
2. the Solow–Swan neoclassical model; and
3. the Romer–Lucas-inspired endogenous growth models.

In each case the ideas developed represent interesting examples of multiple discovery. The first wave of interest focused on the neo-Keynesian work of Roy Harrod (1939, 1948) and Evsey Domar (1946, 1947). In the mid-1950s the development of the neoclassical growth model by Robert Solow (1956) and Trevor Swan (1956) stimulated a second, more lasting and substantial, wave of interest, which, after a period of relative neglect between 1970 and 1986, has been reignited (Mankiw et al., 1992; Mankiw, 1995; Klenow and Rodriguez-Clare, 1997a, 1997b). The third and most recent wave, initiated by the research of Paul Romer (1986) and Robert Lucas (1988), led to the development of endogenous growth theory, which emerged in response to perceived theoretical and empirical deficiencies associated with the neoclassical model (P. Romer, 1994a; Crafts, 1996; Blaug, 2002).

11.9 The Harrod–Domar Model

Following the publication of Keynes's *General Theory* in 1936, some economists sought to dynamize Keynes's static short-run theory in order to investigate the long-run dynamics of capitalist market economies. Roy Harrod (1939, 1948) and Evsey Domar (1946, 1947) independently developed theories that relate an economy's rate of growth to its capital stock. While Keynes emphasized the impact of investment on aggregate demand, Harrod and Domar emphasized how investment spending also increased an economy's productive capacity (a supply-side effect). While Harrod's theory is more ambitious than Domar's, building on Keynesian short-run macroeconomics in order to identify the necessary conditions for equilibrium in a dynamic setting, hereafter we will refer only to the 'Harrod–Domar model', ignoring the subtle differences between the respective contributions of these two outstanding economists.

A major strength of the Harrod–Domar model is its simplicity. The model assumes an exogenous rate of labour force growth (n), a given technology

exhibiting fixed factor proportions (constant capital–labour ratio, K/L) and a fixed capital–output ratio (K/Y). Assuming a two-sector economy (households and firms), we can write the simple national income equation as (11.5):

$$Y_t = C_t + S_t \tag{11.5}$$

where $Y_t = GDP$, C_t = consumption and S_t = saving.

Equilibrium in this simple economy requires (11.6):

$$I_t = S_t \tag{11.6}$$

Substituting (11.6) into (11.5) yields (11.7):

$$Y_t = C_t + I_t \tag{11.7}$$

Within the Harrod–Domar framework the growth of real GDP is assumed to be proportional to the share of investment spending (I) in GDP and for an economy to grow, net additions to the capital stock are required. The evolution of the capital stock over time is given in equation (11.8):

$$K_{t+1} = (1-\delta)K_t + I_t \tag{11.8}$$

where δ is the rate of depreciation of the capital stock. The relationship between the size of the total capital stock (K) and total GDP (Y) is known as the capital–output ratio ($K/Y = v$) and is assumed fixed. Given that we have defined $v = K/Y$, it also follows that $v = \Delta K/\Delta Y$ (where $\Delta K/\Delta Y$ is the incremental capital–output ratio, or ICOR). If we assume that total new investment is determined by total savings, then the essence of the Harrod–Domar model can be set out as follows. Assume that total saving is some proportion (s) of GDP (Y), as shown in equation (11.9):

$$S_t = sY_t \tag{11.9}$$

Since $K = vY$ and $I_t = S_t$, it follows that we can rewrite equation (11.8) as equation (11.10):

$$vY_{t+1} = (1-\delta)vY_t + sY_t \tag{11.10}$$

Dividing through by v, simplifying, and subtracting Y_t from both sides of equation (11.10) yields equation (11.11):

$$Y_{t+1} - Y_t = [s/v - \delta]Y_t \tag{11.11}$$

Dividing through by Y_t gives us equation (11.12):

$$[Y_{t+1} - Y_t]/Y_t = (s/v) - \delta \tag{11.12}$$

Here $[Y_{t+1} - Y_t]/Y_t$ is the growth rate of GDP. Letting $G = [Y_{t+1} - Y_t]/Y_t$, we can write the Harrod–Domar growth equation as (11.13):

$$G = s/v - \delta \tag{11.13}$$

This simply states that the growth rate (G) of GDP is jointly determined by the savings ratio (s) divided by the capital–output ratio (v). The higher the savings ratio and the lower the capital–output ratio and depreciation rate, the faster will an economy grow. In the discussion that follows we will ignore the depreciation rate and consider the Harrod–Domar model as being represented by the equation (11.14):

$$G = s/v \tag{11.14}$$

Thus it is evident from (11.14) that the Harrod–Domar model 'sanctioned the overriding importance of capital accumulation in the quest for enhanced growth' (Shaw, 1992).

The Harrod–Domar model, as Bhagwati recalls, became tremendously influential in the development economics literature during the third quarter of the twentieth century, and was a key component within the framework of economic planning. 'The implications of this popular model were dramatic and reassuring. It suggested that the central developmental problem was simply to increase resources devoted to investment' (Bhagwati, 1984). For example, if a developing country desired to achieve a growth rate of per capita income of 2 per cent per annum (that is, living standards double every 35 years), and population is estimated to be growing at 2 per cent, then economic planners would need to set a target rate of GDP growth (G^*) equal to 4 per cent. If $v = 4$, this implies that G^* can only be achieved with a desired savings ratio (s^*) of 0.16, or 16 per cent of GDP. If $s^* > s$, there is a 'savings gap', and planners needed to devise policies for plugging this gap.

Since the rate of growth in the Harrod–Domar model is positively related to the savings ratio, development economists during the 1950s concentrated their research effort on understanding how to raise private savings ratios in order to enable less developed economies to 'take off' into 'self-sustained growth' (Lewis, 1954, 1955; Rostow, 1960; Easterly, 1999). Reflecting the contemporary development ideas of the 1950s, government fiscal policy was also seen to have a prominent role to play since budgetary surpluses could (in theory) substitute for private domestic savings. If domestic sources of finance

were inadequate to achieve the desired growth target, then foreign aid could fill the 'savings gap' (Riddell, 1987). Aid requirements (Ar) would simply be calculated as $s^* - s = Ar$ (Chenery and Strout, 1966). However, a major weakness of the Harrod–Domar approach is the assumption of a fixed capital–output ratio. Since the inverse of v ($1/v$) is the productivity of investment (ϕ), we can rewrite equation (11.14) as follows:

$$G = s\phi \tag{11.15}$$

Unfortunately, as Bhagwati (1993) observes, the productivity of investment is not a given, but reflects the efficiency of the policy framework and the incentive structures within which investment decisions are taken. The weak growth performance of India before the 1980s reflects, 'not a disappointing savings performance, but rather a disappointing productivity performance' (Bhagwati, 1993). Hence the growth–investment relationship turned out to be 'loose and unstable' due to the multiple factors that influence growth (Easterly, 2001a). Furthermore, economists soon became aware of a second major flaw in the 'aid requirements' or 'financing gap' model. The model assumed that aid inflows would go into investment one to one. But it soon became apparent that inflows of foreign aid, with the objective of closing the savings gap, did not necessarily boost total savings. Aid does not go into investment one to one. Indeed, in many cases inflows of aid led to a reduction of domestic savings together with a decline in the productivity of investment (Griffin, 1970; White, 1992). The research of Boone (1996) confirms that inflows of foreign aid have not raised growth rates in most recipient developing countries. A further problem is that in many developing countries the 'soft budget constraints' operating within the public sector created a climate for what Bhagwati calls 'goofing off'. It is therefore hardly surprising that public sector enterprises frequently failed to generate profits intended to add to government saving. In short, 'capital fundamentalism' and the 'aid-financed investment fetish', which dominated development thinking for much of the period after 1950, led economists up the wrong path in their 'elusive quest for growth' (King and Levine, 1994; Easterly, 2001a, 2003; Easterly et al., 2003; Snowdon, 2003a). Indeed, William Easterly (1999), a former World Bank economist, argues that the Harrod–Domar model is far from dead and still continues to exercise considerable influence on economists working within the major international financial institutions even if it died long ago in the academic literature. Easterly shows that economists working at the World Bank, International Monetary Fund, Inter-American Bank, European Bank for Reconstruction and Development, and the International Labour Organization still frequently employ the Harrod–Domar–Chenery–Strout methodology to calculate the investment and aid requirements needed in order for specific

countries to achieve their growth targets. However, as Easterly convincingly demonstrates, the evidence that aid flows into investment on a one-for-one basis, and that there is a fixed linear relationship between growth and investment in the short run, is 'soundly rejected'.

A further weakness of the Harrod–Domar framework is the assumption of zero substitutability between capital and labour (that is, a fixed factor proportions production function). This is a 'crucial' but inappropriate assumption for a model concerned with long-run growth. This assumption of the Harrod–Domar model also leads to the renowned instability property that 'even for the long run an economic system is at best balanced on a knife-edge equilibrium growth' (Solow, 1956). In Harrod's model the possibility of achieving steady growth with full employment was remote. Only in very special circumstances will an economy remain in equilibrium with full employment of both labour and capital. As Solow (1988) noted in his Nobel Memorial lecture, to achieve steady growth in a Harrod–Domar world would be 'a miraculous stroke of luck'. The problem arises from the assumption of a production function with an inflexible technology. In the Harrod–Domar model the capital–output ratio (K/Y) and the capital–labour ratio (K/L) are assumed constant. In a growth setting this means that K and Y must always grow at the same rate to maintain equilibrium. However, because the model also assumes a constant capital–labour ratio (K/L), K and L must also grow at the same rate. Therefore, if we assume that the labour force (L) grows at the same rate as the rate of growth of population (n), then we can conclude that the only way that equilibrium can be maintained in the model is for $n = G = s/v$. It would only be by pure coincidence that $n = G$. If $n > G$, the result will be continually rising unemployment. If $G > n$, the capital stock will become increasingly idle and the growth rate of output will slow down to $G = n$. Thus, whenever K and L do not grow at the same rate, the economy falls off its equilibrium 'knife-edge' growth path. However, the evidence is overwhelming that this property does not fit well with the actual experience of growth (for a more detailed discussion of the Harrod–Domar model see Hahn and Matthews, 1964; H. Jones, 1975).

11.10 The Solow Neoclassical Growth Model

Following the seminal contributions of Solow (1956, 1957) and Swan (1956), the neoclassical model became the dominant approach to the analysis of growth, at least within academia. Between 1956 and 1970 economists refined 'old growth theory', better known as the Solow neoclassical model of economic growth (Solow, 2000, 2002). Building on a neoclassical production function framework, the Solow model highlights the impact on growth of saving, population growth and technolgical progress in a closed economy

setting without a government sector. Despite recent developments in endogenous growth theory, the Solow model remains the essential starting point to any discussion of economic growth. As Mankiw (1995, 2003) notes, whenever practical macroeconomists have to answer questions about long-run growth they usually begin with a simple neoclassical growth model (see also Abel and Bernanke, 2001; Jones, 2001a; Barro and Sala-i-Martin, 2003).

The key assumptions of the Solow model are: (i) for simplicity it is assumed that the economy consists of one sector producing one type of commodity that can be used for either investment or consumption purposes; (ii) the economy is closed to international transactions and the government sector is ignored; (iii) all output that is saved is invested; that is, in the Solow model the absence of a separate investment function implies that Keynesian difficulties are eliminated since *ex ante* saving and *ex ante* investment are always equivalent; (iv) since the model is concerned with the long run there are no Keynesian stability problems; that is, the assumptions of full price flexibility and monetary neutrality apply and the economy is always producing its potential (natural) level of total output; (v) Solow abandons the Harrod–Domar assumptions of a fixed capital–output ratio (K/Y) and fixed capital–labour ratio (K/L); (vi) the rate of technological progress, population growth and the depreciation rate of the capital stock are all determined exogenously.

Given these assumptions we can concentrate on developing the three key relationships in the Solow model, namely, the production function, the consumption function and the capital accumulation process.

The production function

The Solow growth model is built around the neoclassical aggregate production function (11.16) and focuses on the *proximate* causes of growth:

$$Y = A_t F(K, L) \tag{11.16}$$

where Y is real output, K is capital, L is the labour input and A_t is a measure of technology (that is, the way that inputs to the production function can be transformed into output) which is exogenous and taken simply to depend on time. Sometimes, A_t is called 'total factor productivity'. It is important to be clear about what the assumption of exogenous technology means in the Solow model. In the neoclassical theory of growth, technology is assumed to be a public good. Applied to the world economy this means that every country is assumed to share the same stock of knowledge which is freely available; that is, all countries have access to the same production function. In his defence of the neoclassical assumption of treating technology as if it were a public good, Mankiw (1995) puts his case as follows:

> The production function should not be viewed literally as a description of a specific production process, but as a mapping from quantities of inputs into a quantity of output. To say that different countries have the same production function is merely to say that if they had the same inputs, they would produce the same output. Different countries with different levels of inputs need not rely on exactly the same processes for producing goods and services. When a country doubles its capital stock, it does not give each worker twice as many shovels. Instead, it replaces shovels with bulldozers. For the purposes of modelling economic growth, this change should be viewed as a movement along the same production function, rather than a shift to a completely new production function.

As we shall see later (section 11.15), many economists disagree with this approach and insist that there are significant technology gaps between nations (see Fagerberg, 1994; P. Romer, 1995). However, to progress with our examination of the Solow model we will continue to treat technology as a public good.

For simplicity, let us begin by first assuming a situation where there is no technological progress. Making this assumption of a given state of technology will allow us to concentrate on the relationship between output per worker and capital per worker. We can therefore rewrite (11.16) as:

$$Y = F(K, L) \tag{11.17}$$

The aggregate production function given by (11.17) is assumed to be 'well behaved'; that is, it satisfies the following three conditions (see Inada, 1963; D. Romer, 2001; Barro and Sala-i-Martin, 2003; Mankiw, 2003). First, for all values of $K > 0$ and $L > 0$, $F(\cdot)$ exhibits positive but diminishing marginal returns with respect to both capital and labour; that is, $\partial F/\partial K > 0$, $\partial^2 F/\partial K^2 < 0$, $\partial F/\partial L > 0$, and $\partial^2 F/\partial L^2 < 0$. Second, the production function exhibits constant returns to scale such that $F(\lambda K, \lambda L) = \lambda Y$; that is, raising inputs by λ will also increase aggregate output by λ. Letting $\lambda = 1/L$ yields $Y/L = F(K/L)$. This assumption allows (11.17) to be written down in intensive form as (11.18), where y = output per worker (Y/L) and k = capital per worker (K/L):

$$y = f(k), \text{ where } f'(k) > 0, \text{ and } f''(k) < 0 \text{ for all } k \tag{11.18}$$

Equation (11.18) states that output per worker is a positive function of the capital–labour ratio and exhibits diminishing returns. The key assumption of constant returns to scale implies that the economy is sufficiently large that any Smithian gains from further division of labour and specialization have already been exhausted, so that the size of the economy, in terms of the labour force, has no influence on output per worker. Third, as the capital–labour ratio approaches infinity ($k \rightarrow \infty$) the marginal product of capital (*MPK*)

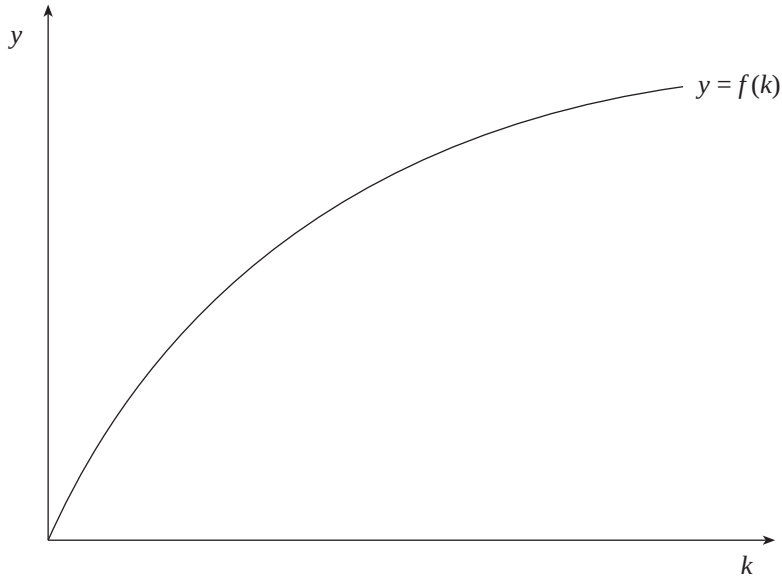


Figure 11.3 The neoclassical aggregate production function

approaches zero; as the capital–labour ratio approaches zero the marginal product of capital tends towards infinity ($MPK \rightarrow \infty$).

Figure 11.3 shows an intensive form of the neoclassical aggregate production function that satisfies the above conditions. As the diagram illustrates, for a given technology, any country that increases its capital–labour ratio (more equipment per worker) will have a higher output per worker. However, because of diminishing returns, the impact on output per worker resulting from capital accumulation per worker (capital deepening) will continuously decline. Thus for a given increase in k, the impact on y will be much greater where capital is relatively scarce than in economies where capital is relatively abundant. That is, the accumulation of capital should have a much more dramatic impact on labour productivity in developing countries compared to developed countries.

The slope of the production function measures the marginal product of capital, where $MPK = f(k + 1) - f(k)$. In the Solow model the *MPK* should be much higher in developing economies compared to developed economies. In an open economy setting with no restrictions on capital mobility, we should therefore expect to see, *ceteris paribus*, capital flowing from rich to poor countries, attracted by higher potential returns, thereby accelerating the process of capital accumulation.

The consumption function

Since output per worker depends positively on capital per worker, we need to understand how the capital–labour ratio evolves over time. To examine the process of capital accumulation we first need to specify the determination of saving. In a closed economy aggregate output = aggregate income and comprises two components, namely, consumption (C) and investment (I) = Savings (S). Therefore we can write equation (11.19) for income as:

$$Y = C + I \tag{11.19}$$

or equivalently

$$Y = C + S$$

Here $S = sY$ is a simple savings function where s is the fraction of income saved and $1 > s > 0$. We can rewrite (11.19) as (11.20):

$$Y = C + sY \tag{11.20}$$

Given the assumption of a closed economy, private domestic saving (sY) must equal domestic investment (I).

The capital accumulation process

A country's capital stock (K_t) at a point in time consists of plant, machinery and infrastructure. Each year a proportion of the capital stock wears out. The parameter δ represents this process of depreciation. Countering this tendency for the capital stock to decline is a flow of investment spending each year (I_t) that adds to the capital stock. Therefore, given these two forces, we can write an equation for the evolution of the capital stock of the following form:

$$K_{t+1} = I_t + (1-\delta)K_t = sY_t + K_t - \delta K_t \tag{11.21}$$

Rewriting (11.21) in per worker terms yields equation (11.22):

$$K_{t+1}/L = sY_t/L + K_t/L - \delta K_t/L \tag{11.22}$$

Deducting K_t/L from both sides of (11.22) gives us (11.23):

$$K_{t+1}/L - K_t/L = sY_t/L - \delta K_t/L \tag{11.23}$$

In the neoclassical theory of growth the accumulation of capital evolves according to (11.24), which is the fundamental differential equation of the Solow model:

$$\dot{k} = sf(k) - \delta k \tag{11.24}$$

where $\dot{k} = K_{t+1}/L - K_t/L$ is the change of the capital input per worker, and $sf(k) = sy = sY_t/L$ is saving (investment) per worker. The $\delta k = \delta K_t/L$ term represents the 'investment requirements' per worker in order to keep the capital–labour ratio constant. The steady-state condition in the Solow model is given in equation (11.25):

$$sf(k^*) - \delta k^* = 0 \tag{11.25}$$

Thus, in the steady state $sf(k^*) = \delta k^*$; that is, investment per worker is just sufficient to cover depreciation per worker, leaving capital per worker constant.

Extending the model to allow for growth of the labour force is relatively straightforward. In the Solow model it is assumed that the participation rate is constant, so that the labour force grows at a constant proportionate rate equal to the exogenously determined rate of growth of population = n. Because $k = K/L$, population growth, by increasing the supply of labour, will reduce k. Therefore population growth has the same impact on k as depreciation. We need to modify (11.24) to reflect the influence of population growth. The fundamental differential equation now becomes:

$$\dot{k} = sf(k) - (n + \delta)k \tag{11.26}$$

We can think of the expression $(n + \delta)k$ as the 'required' or 'break-even' investment necessary to keep the capital stock per unit of labour (k) constant. In order to prevent k from falling, some investment is required to offset depreciation. This is the $(\delta)k$ term in (11.26). Some investment is also required because the quantity of labour is growing at a rate = n. This is the $(n)k$ term in (11.26). Hence the capital stock must grow at rate $(n + \delta)$ just to hold k steady. When investment per unit of labour is greater than required for break-even investment, then k will be rising and in this case the economy is experiencing 'capital deepening'. Given the structure of the Solow model the economy will, in time, approach a steady state where actual investment per worker, $sf(k)$, equals break-even investment per worker, $(n + \delta)k$. In the steady state the change in capital per worker $\dot{k} = 0$, although the economy continues to experience 'capital widening', the extension of existing capital per worker to additional workers. Using * to indicate steady-state values, we can define the steady state as (11.27):

$$sf(k^*) = (n + \delta)k^* \tag{11.27}$$

Figure 11.4 captures the essential features of the Solow model outlined by equations (11.18) to (11.27). In the top panel of Figure 11.4 the curve $f(k)$

graphs a well-behaved intensive production function; $sf(k)$ shows the level of savings per worker at different levels of the capital–labour ratio (k); the linear relationship $(n + \delta)k$ shows that break-even investment is proportional to k. At the capital–labour ratio k_1, savings (investment) per worker (b) exceed required investment (c) and so the economy experiences capital deepening and k rises. At k_1 consumption per worker is indicated by $d - b$ and output per worker is y_1. At k_2, because $(n + \delta)k > sf(k)$ the capital–labour ratio falls, capital becomes 'shallower' (Jones, 1975). The steady state balanced growth path occurs at k^*, where investment per worker equals break-even investment. Output per worker is y^* and consumption per worker is $e - a$. In the bottom panel of Figure 11.4 the relationship between $\dot{k}$ (the change of the capital–labour ratio) and k is shown with a phase diagram. When $\dot{k} > 0$, k is rising; when $\dot{k} < 0$, k is falling.

In the steady state equilibrium, shown as point a in the top panel of Figure 11.4, output per worker (y^*) and capital per worker (k^*) are constant. However, although there is no intensive growth in the steady state, there is extensive growth because population (and hence the labour input = L) is growing at a rate of n per cent per annum. Thus, in order for $y^* = Y/L$ and $k^* = K/L$ to remain constant, both Y and K must also grow at the same rate as population.

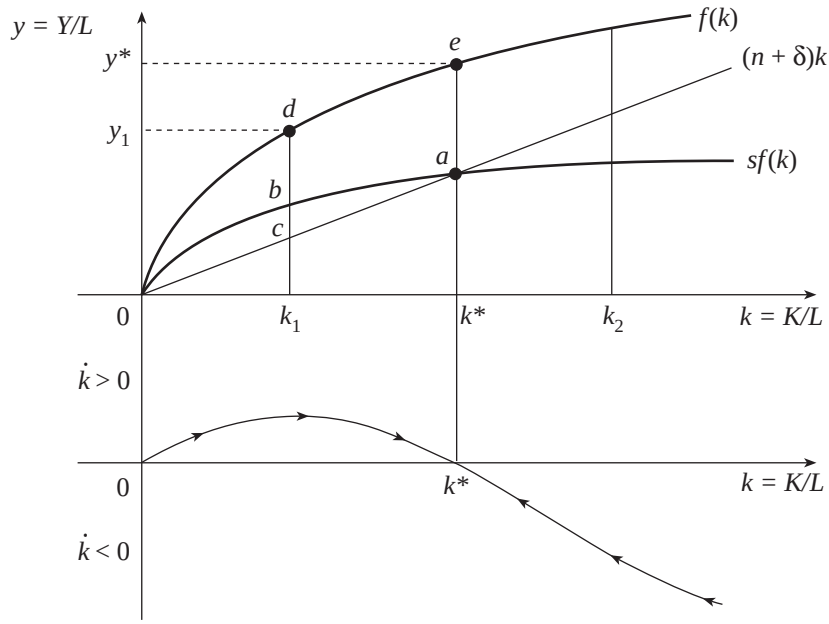


Figure 11.4 The Solow growth model

It can be seen from Figure 11.4 that the steady state level of output per worker will increase (*ceteris paribus*) if the rate of population growth and/or the depreciation rate are reduced (a downward pivot of the $(n + \delta)k$ function), and vice versa. The steady state level of output per worker will also increase (*ceteris paribus*) if the savings rate increases (an upward shift of the *sf(k)* function), and vice versa. Of particular importance is the prediction from the Solow model that an increase in the savings ratio cannot permanently increase the long-run rate of growth. A higher savings ratio does temporarily increase the growth rate during the period of transitional dynamics to the new steady state and it also permanently increases the level of output per worker. Of course the period of transitional dynamics may be a long historical time period and level effects are important and should not be undervalued (see Solow, 2000; Temple, 2003).

So far we have assumed zero technological progress. Given the fact that output per worker has shown a continuous tendency to increase, at least since the onset of the Industrial Revolution in the now developed economies, a model that predicts a constant steady state output per worker is clearly unsatisfactory. A surprising conclusion of the neoclassical growth model is that without technological progress the ability of an economy to raise output per worker via capital accumulation is limited by the interaction of diminishing returns, the willingness of people to save, the rate of population growth, and the rate of depreciation of the capital stock. In order to explain continuous growth of output per worker in the long run the Solow model must incorporate the influence of sustained technological progress.

The production function (11.16), in its Cobb–Douglas form, can be written as (11.28):

$$Y = A_t K^{\alpha} L^{1-\alpha} \tag{11.28}$$

where α and $1 - \alpha$ are weights reflecting the share of capital and labour in the national income. Assuming constant returns to scale, output per worker (Y/L) is not affected by the scale of output, and, for a given technology, A_{t0}, output per worker is positively related to the capital–labour ratio (K/L). We can therefore rewrite the production function equation (11.28) in terms of output per worker as shown by equation (11.29):

$$Y/L = A(t_0)(K/L) = A(t_0)K^{\alpha}L^{1-\alpha}/L = A(t_0)(K/L)^{\alpha} \tag{11.29}$$

Letting $y = Y/L$ and $k = K/L$, we finally arrive at the 'intensive form' of the aggregate production function shown in equation (11.30):

$$y = A(t_0)k^{\alpha} \tag{11.30}$$

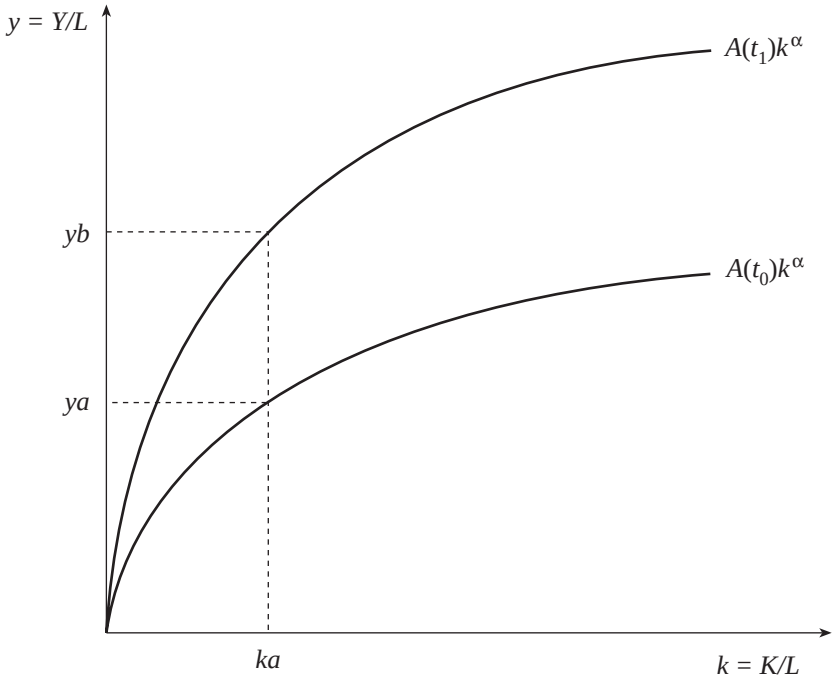


Figure 11.5 Technological progress

For a given technology, equation (11.30) tells us that increasing the amount of capital per worker (capital deepening) will lead to an increase in output per worker. The impact of exogenous technological progress is illustrated in Figure 11.5 by a shift of the production function between two time periods ($t_0 \Rightarrow t_1$) from $A(t_0)k^{\alpha}$ to $A(t_1)k^{\alpha}$, raising output per worker from *ya* to *yb* for a given capital–labour ratio of *ka*. Continuous upward shifts of the production function, induced by an exogenously determined growth of knowledge, provide the only mechanism for 'explaining' steady state growth of output per worker in the neoclassical model.

Therefore, although it was not Solow's original intention, it was his neoclassical theory of growth that brought technological progress to prominence as a major explanatory factor in the analysis of economic growth. But, somewhat paradoxically, in Solow's theory technological progress is exogenous, that is, not explained by the model! Solow admits that he made technological progress exogenous in his model in order to simplify it and also because he did not 'pretend to understand' it (see Solow interview at the end of this chapter) and, as Abramovitz (1956) observed, the Solow residual turned out to be 'a measure of our ignorance' (see also Abramovitz, 1999). While Barro

and Sala-i-Martin (1995) conclude that this was 'an obviously unsatisfactory situation', David Romer (1996) comments that the Solow model 'takes as given the behaviour of the variable that it identifies as the main driving force of growth'. Furthermore, although the Solow model attributes no role to capital accumulation in achieving long-run sustainable growth, it should be noted that productivity growth may not be independent of capital accumulation if technical progress is embodied in new capital equipment. Unlike disembodied technical progress, which can raise the productivity of the existing inputs, embodied technical progress does not benefit older capital equipment. It should also be noted that DeLong and Summers (1991, 1992) find a strong association between equipment investment and economic growth in the period 1960–85 for a sample of over 60 countries.

Remarkably, while economists have long recognized the crucial importance of technological change as a major source of dynamism in capitalist economies (especially Karl Marx and Joseph Schumpeter), the analysis of technological change and innovation by economists has, until recently, been an area of relative neglect (see Freeman, 1994; Baumol, 2002).

Leaving aside these controversies for the moment, it is important to note that the Solow model allows us to make several important predictions about the growth process (see Mankiw, 1995, 2003; Solow, 2002):

1. in the long run an economy will gradually approach a steady state equilibrium with y^* and k^* independent of initial conditions;
2. the steady state balanced rate of growth of aggregate output depends on the rate of population growth (n) and the rate of technological progress (A);
3. in the steady state balanced growth path the rate of growth of output per worker depends solely on the rate of technological progress. As illustrated in Figure 11.5, without technological progress the growth of output per worker will eventually cease;
4. the steady state rate of growth of the capital stock equals the rate of income growth, so the K/Y ratio is constant;
5. for a given depreciation rate (δ) the steady state level of output per worker depends on the savings rate (s) and the population growth rate (n). A higher rate of saving will increase y^*, a higher population growth rate will reduce y^*;
6. the impact of an increase in the savings (investment) rate on the growth of output per worker is temporary. An economy experiences a period of higher growth as the new steady state is approached. A higher rate of saving has no effect on the long-run sustainable rate of growth, although it will increase the level of output per worker. To Solow this finding was a 'real shocker';

7. the Solow model has particular 'convergence properties'. In particular, 'if countries are similar with respect to structural parameters for preferences and technology, then poor countries tend to grow faster than rich countries' (Barro, 1991).

The result in the Solow model that an increase in the saving rate has no impact on the long-run rate of economic growth contains 'more than a touch of irony' (Cesaratto, 1999). As Hamberg (1971) pointed out, the neo-Keynesian Harrod–Domar model highlights the importance of increasing the saving rate to increase long-run growth, while in Keynes's (1936) *General Theory* an increase in the saving rate leads to a fall in output in the short run through its negative impact on aggregate demand (the so-called 'paradox of thrift' effect). In contrast, the long tradition within classical–neoclassical economics of highlighting the virtues of thrift come a little unstuck with the Solow model since it is technological progress, not thrift, that drives long-run growth of output per worker (see Cesaratto, 1999)!

11.11 Accounting for the Sources of Economic Growth

Economists not only need a theoretical framework for understanding the causes of growth; they also require a simple method of calculating the relative importance of capital, labour and technology in the growth experience of actual economies. The established framework, following Solow's (1957) seminal contribution, is called 'growth accounting' (see Abel and Bernanke, 2001. Some economists remain highly sceptical about the whole methodology and theoretical basis of growth accounting, for example Nelson, 1973). As far as the proximate causes of growth are concerned we can see by referring back to equation (11.28) that increases in total GDP (Y) come from the combined weighted impact of capital accumulation, labour supply growth and technological progress. Economists can measure changes in the amount of capital and labour that occur in an economy over time, but changes in technology (total factor productivity = TFP) are not directly observable. However, it is possible to measure changes in TFP as a 'residual' after taking into account the contributions to growth made by changes in the capital and labour inputs. Solow's (1957) technique was to define technological change as changes in aggregate output minus the sum of the weighted contributions of the labour and capital inputs. In short, the Solow residual measures that part of a change in aggregate output which cannot be explained by changes in the measurable quantities of capital and labour inputs. The derivation of the Solow residual can be shown as follows. The aggregate production function in equation (11.28) shows that output (Y) is dependent on the inputs of capital (K), labour (L) and the currently available technology (A), which acts as an index of total

factor productivity. Output will change if *A*, *K* or *L* change. In equation (11.28) the exponent on the capital shock α measures the elasticity of output with respect to capital and the exponent on the labour input $(1 - \alpha)$ measures the elasticity of output with respect to labour. The weights α and $1 - \alpha$ are estimated from national income statistics and reflect the income shares of capital and labour respectively. Since these weights sum to unity, this indicates that (11.28) is a constant returns to scale production function. Hence an equal percentage increase in both factor inputs (*K* and *L*) will increase *Y* by the same percentage. Since the growth rate of the product of the inputs will be the growth rate of *A* plus the growth rate of K^{α} plus the growth rate of $L^{1-\alpha}$, equation (11.28) can be rewritten as (11.31), which is the basic growth accounting equation used in numerous empirical studies of the sources of economic growth (see Maddison, 1972, 1987; Denison, 1985; Young, 1995, Crafts, 2000; Jorgenson, 2001).

$$\Delta Y/Y = \Delta A/A + \alpha\Delta K/K + (1-\alpha)\Delta L/L \tag{11.31}$$

Equation (11.31) is simply the Cobb–Douglas production function written in a form representing rates of change. It shows that the growth of aggregate output $(\Delta Y/Y)$ depends on the contribution of changes in total factor productivity $(\Delta A/A)$, changes in the weighted contribution of capital, $\alpha\Delta K/K$, and changes in the weighted contribution of labour $(1 - \alpha)\Delta L/L$. By rearranging equation (11.28) we can represent the productivity index (TFP) which we need to measure as equation (11.32):

$$TFP = A = Y / K^{\alpha} L^{1-\alpha} \tag{11.32}$$

As already noted, because there is no direct way of measuring TFP it has to be estimated as a residual. By writing down equation (11.32) in terms of rates of change we can obtain an equation from which the growth of TFP (technological change) can be estimated as a residual. This is shown in equation (11.33):

$$\Delta A/A = \Delta Y/Y - [\alpha\Delta K/K + (1-\alpha)\Delta L/L] \tag{11.33}$$

Data relating to output and the capital and labour inputs are available. Estimates of α and hence $1 - \alpha$ can be acquired from historical national income data. For example, in Solow's original paper covering the US economy for the period 1909–49 he estimated that the rate of growth of total output $(\Delta Y/Y)$ had averaged 2.9 per cent per year, of which 0.32 percentage points could be attributed to capital $(\alpha\Delta K/K)$, 1.09 percentage points could be attributed to labour $(1 - \alpha\Delta L/L)$, leaving a 'Solow residual' $(\Delta A/A)$ of 1.49 percentage

points. In other words, almost half of the growth experienced in the USA during this period was due to unexplained technological progress! In Denison's (1985) later work he found that for the period 1929–82, $\Delta Y/Y = 2.92$ per cent, of which 1.02 percentage points were be attributed to $\Delta A/A$. More recent controversial research by Alwyn Young (1992, 1994, 1995) on the sources of growth in the East Asian Tiger economies has suggested estimates of rates of growth of TFP for Taiwan of 2.6 per cent, for South Korea of 1.7 per cent, for Hong Kong of 1.7 per cent and for Singapore a meagre 0.2 per cent! So although these economies have experienced unprecedented growth rates of GDP since the early 1960s, Young's research suggests that these economies are examples of miracles of accumulation. Once we account for the growth of labour and physical and human capital there is little left to explain, especially in the case of Singapore (see Krugman, 1994b; Hsieh, 1999; Bhagwati, 2000). Going further back in history, Nick Crafts (1994, 1995) has provided estimates of the sources of growth for the British economy during the period 1760–1913. Crafts's estimates suggest that 'by twentieth century standards both the output growth rates and the TFP rates are quite modest' (Crafts, 1995).

The most obvious feature of the post-1973 growth accounting data is the well-known puzzle of the 'productivity slowdown'. This slowdown has been attributed to many possible causes, including the adverse impact on investment and existing capital stocks of the 1970s oil price shocks, a slowdown in the rate of innovation, adverse demographic trends, an increasingly regulatory environment and problems associated with measurement such as accounting for quality changes (Fischer et al., 1988).

In a recent survey of the growth accounting literature Bosworth and Collins (2003) reaffirm their belief that growth accounting techniques can yield useful and consistent results. In the debate over the relative importance of capital accumulation v. TFP in accounting for growth Bosworth and Collins conclude that 'both are important' and that 'some of the earlier research understates the role of capital accumulation because of inadequate measurement of the capital input'.

11.12 The Convergence Debate

Since 1945 the economies of what used to known as the Third World have been viewed as participating in an attempt to achieve economic development and thereby begin to 'catch up' the rich countries of the world in terms of per capita income. The growing awareness of the wide variety of experiences observed among developing countries in this attempt has been a major factor in motivating renewed research into the important issue of economic growth.

> It is generally accepted that the Third World's efforts to join the ranks of the so-called 'mature industrial countries' represent one of the major social, economic and political phenomena of the second half of the twentieth century. This attempted transition to modern economic growth will rank with the taming of the atom as the most important event of this period. (Fei and Ranis, 1997)

Modern discussion of the convergence issue began with the contribution of Gerschenkron (1962), who argued that poor countries could benefit from the advantages of 'relative backwardness' since the possibilities of technological transfer from the developed countries could vastly speed up the pace of industrialization. However, this debate has much earlier origins, dating back to 1750, when Hume put forward the view that the growth process would eventually generate convergence because economic growth in the rich countries would exhibit a natural tendency to slow through a process of 'endogenous decay' (Elmslie and Criss, 1999). Oswald and Tucker (see Elmslie and Criss, 1999) rejected Hume's arguments, putting forward an endogenous growth view that 'increasing, or at least non-decreasing, returns in both scientific and economic activity will keep poor countries from naturally converging towards their rich neighbours'. Elsewhere, Elmslie has also argued that in the *Wealth of Nations*, Smith (1776) took up an endogenous growth position since societal extensions to the division of labour will allow the rich countries to continuously maintain or extend their technological lead over poorer countries (see Elmslie and Criss, 1999). This argument also lies at the heart of Babbage's 1835 thesis that the perpetual advances in science provide the foundation for further advancement and economic progress. Elmslie and Criss argue that Babbage's case against the restrictive laws on the export of machines is 'the best statement of endogenous growth in the classical period'. For, as Babbage argued, the growth of other countries does not pose an economic threat because 'the sun of science has yet penetrated but through the outer fold of Nature's majestic robe'.

In more recent times the issue of convergence began to receive a great deal of attention from the mid-1980s and this growth of research interest stems mainly from the growing recognition that many poor economies were failing to exhibit a tendency to close the per capita incomes gap with rich countries (see Islam, 2003). The conundrum of non-convergence of per capita incomes across the world's economies was first clearly articulated by Paul Romer (1986). The convergence property in the Solow model stems from the key assumption of diminishing returns to reproducible capital. With constant returns to scale, a proportional increase in the inputs of labour and capital leads to a proportional increase in output. By increasing the capital–labour ratio an economy will experience diminishing marginal productivity of capital. Hence poor countries with low capital-to-labour ratios have high marginal products of capital and consequently high growth rates for a given rate of

investment. In contrast, rich countries have high capital-to-labour ratios, low marginal products of capital and hence low growth rates (see the aggregate production function $A(t_0)k^{\alpha}$ in Figure 11.5). The severity of diminishing returns depends on the relative importance of capital in the production process and hence the size of the capital share (α) determines the curvature of the production function and the speed at which diminishing returns set in (see DeLong, 2001). With a small capital share (typically $\alpha = 1/3$), the average and marginal product of labour declines rapidly as capital deepening takes place. It is obvious from an inspection of the production function in Figures 11.3–11.5 that in the Solow model capital accumulation has a much bigger impact on output per worker when capita per worker ratios are low compared to when they are high. In a risk-free world with international capital mobility this tendency for convergence will be reinforced (Lucas, 1990b). In the long run the neoclassical model also predicts convergence of growth rates for economies which have reached their steady state. However, as pointed out by Romer, the neoclassical hypothesis that low income per capita economies will tend to grow faster than high income per capita economies appears to be inconsistent with the cross-country evidence.

In his seminal 1986 paper Romer raised important doubts about the preference economists display for a growth model which exhibits diminishing returns to capital accumulation, falling rates of growth over time, and convergence of per capita income levels and growth rates across countries. Evidence relating to falling rates of growth can be found by examining the historical growth record of 'leader' economies compared to other economies (where leader is defined in terms of the highest level of productivity). Maddison (1982) has identified three leader economies since 1700, namely: the Netherlands, 1700–85; the UK, 1785–1890; and the USA, 1890–1979. As the twenty-first century begins, the USA remains the leader economy. But, as Romer notes, the rate of growth has been increasing for the leader economies from essentially zero in eighteenth-century Netherlands to 2.3 per cent per annum for the USA in the period 1890–1979. Historical data for industrial countries also indicate a positive rather than negative trend for growth rates. Hence, rather than modify the neoclassical growth model, Romer introduced an alternative endogenous theory of growth where there is no steady state level of income, where growth rates can increase over time, and where income per capita differentials between countries can persist indefinitely.

The general property of convergence is often presented as a tendency of poor countries to have higher rates of growth than the average and for rich countries to grow more slowly than average. In the world as a whole 'no such tendency is found' (Sachs and Warner, 1995). However, there is strong evidence of convergence among the OECD economies as well as between US states, Japanese prefectures and European regions within the European Com-

munity (Baumol, 1986; DeLong, 1988; Dowrick, 1992; Barro and Sala-i-Martin, 2003). The conflicting evidence led Baumol to suggest that there may be a 'convergence club' whereby only those countries with an adequate human capital base and favourable institutions can hope to participate in convergent growth. More recently, DeLong and Dowrick (2002) have shown that 'what convergence there has been has been limited in geography and time' and, as a result, to use Pritchett's (1997) words, there has been 'Divergence, Big Time' (see Jones, 1997a, 1997b; Melchior, 2001).

The research inspired by Barro (1991) has shown how the prediction of convergence in the neoclassical model needs considerable qualification. If all economies had identical savings rates, population growth rates and unlimited access to the same technology, then relative capital intensities would determine output per capita differentials between countries. Poor countries with low capital intensities are predicted to grow faster than rich countries in the period of transitional dynamics en route to the common steady state equilibrium. In this situation there will be unconditional or absolute convergence. Clearly, given the restrictive requirements, this outcome is only likely to be observed among a group of relatively homogeneous countries or regions that share similar characteristics, such as the OECD economies and US states. In reality, many economies differ considerably with respect to key variables (such as saving propensities, government policies and population growth) and are moving towards different steady states. Therefore the general convergence property of the Solow model is conditional. 'Each economy converges to its own steady state, which in turn is determined by its saving and population growth rates' (Mankiw, 1995). This property of conditional convergence implies that growth rates will be rapid during transitional dynamics if a country's initial output per capita is low relative to its long-run steady state value. When countries reach their respective steady states, growth rates will then equalize in line with the rate of technological progress. Clearly, if rich countries have higher steady state values of k^* than poor countries, there will be no possibility of convergence in an absolute sense. As Barro (1997) notes, 'a poor country that also has a low long-term position, possibly because its public policies are harmful or its saving rate is low, would not tend to grow rapidly'. Conditional convergence therefore allows for the possibility that rich countries may grow faster than poor countries, leading to income per capita divergence! Since countries do not have the same steady state per capita income, each country will have a tendency to grow more rapidly the bigger the gap between its initial level of income per capita and its own long-run steady state per capita income.

This can be illustrated as follows. Abstracting from technological progress, we have the intensive form of the production function written as (11.34):

$$y = k^{\alpha} \tag{11.34}$$

Expressing (11.34) in terms of growth rates gives (11.35):

$$\dot{y}/y = \alpha \dot{k}/k \tag{11.35}$$

Dividing both sides of Solow's fundamental equation (11.26) by k gives equation (11.36):

$$\dot{k}/k = sf(k)/k - (n + \delta) \tag{11.36}$$

Therefore, substituting (11.35) into (11.36), we derive an expression for the growth rate of output per worker given by equation (11.37):

$$\dot{y}/y = \alpha[sf(k)/k - (n + \delta)] \tag{11.37}$$

In Figure 11.6 the growth rate of the capital–labour ratio $(\dot{k}/k)$ is shown by the vertical distance between the $sf(k)/k$ function and the effective depreciation line, $n + \delta$ (see Jones, 2001a; Barro and Sala-i-Martin, 2003). The intersection of the savings curve and effective depreciation line determines the steady state capital per worker, k^*. In Figure 11.7 we compare a rich

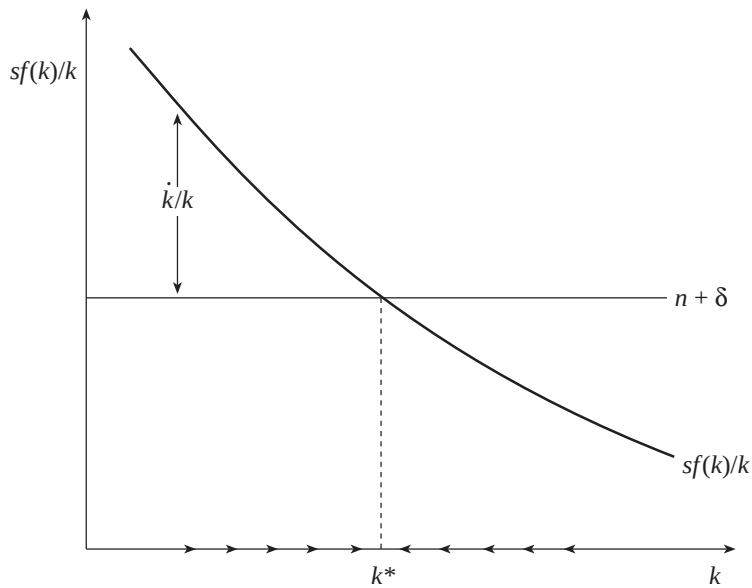


Figure 11.6 Transition dynamics

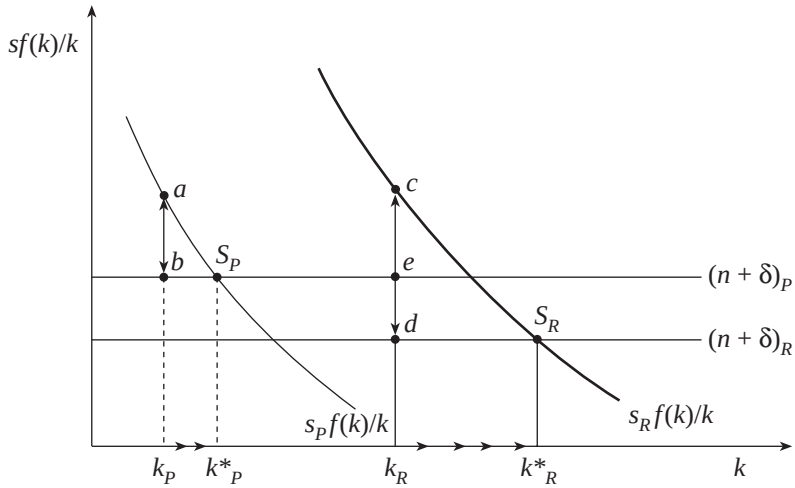


Figure 11.7 Conditional convergence

developed country with a poor developing country. Here we assume (realistically) that the developing country has a higher rate of population growth than the developed country, that is, $(n + \delta)_P > (n + \delta)_R$, and also that the developed country has a higher savings rate than the developing country. The steady state for the developing country is indicated by point S_P, with a steady state capital–labour ratio of k^*_P. Similarly, the steady state for the developed country is indicated by points S_R and k^*_R. Suppose the current location of these economies is given by k_P and k_R. It is clear that the developed economy will be growing faster than the developing country because the rate of growth of the capital–labour ratio is greater in the developed economy (distance *c*–*d*) than the developing country (*a*–*b*). Figure 11.7 also shows that even if the developed country had the same population growth rate as the developing country it would still have a faster rate of growth since the gap between the savings curve and the effective depreciation line is still greater than that for the developing country, that is, $a–b < c–e$.

Robert Lucas (2000b) has recently presented a numerical simulation of world income dynamics in a model which captures certain features of the diffusion of the Industrial Revolution across the world's economies (see Snowdon, 2002a). In discussing prospects for the twenty-first century Lucas concludes from his simulation exercise that 'the restoration of inter-society income equality will be one of the major economic events of the century to come'. In the twenty-first century we will witness 'Convergence, Big Time'! In short, we will witness an ever-growing 'convergence club' as sooner or later 'everyone will join the Industrial Revolution'.

In Lucas's model the followers grow faster than the leader and will eventually converge on the income per capita level of the leader, 'but will never surpass the leader's level'. As followers catch up the leader Lucas assumes that their growth rates converge towards that of the leader, that is, 2 per cent. The probability that a pre-industrial country will begin to grow is positively related to the level of production in the rest of the world which in turn reflects past growth experienced. There are several possible sources of the diffusion of the Industrial Revolution from leaders to followers, for example:

1. diffusion via spillovers due to human capital externalities (Tamura, 1996), the idea that 'knowledge produced anywhere benefits people everywhere';
2. diffusion via adopting the policies and institutions of the successful countries thus removing the barriers to growth (Olson, 1996; Parente and Prescott, 1999, 2000);
3. diffusion due to diminishing returns leading to capital flows to the low-income economies (Lucas, 1990b).

Lucas's simulations predict that the diffusion of the Industrial Revolution was relatively slow for the nineteenth century but accelerated 'dramatically' in the twentieth century, finally slowing down towards the year 2000 'because there are so few people left in stagnant, pre-industrial economies'. In Lucas's simulation, by the year 2000, 90 per cent of the world is growing. Given the rate of diffusion, world income inequality at first increases, peaking some time in the 1970s, and then declines, 'ultimately to zero'. According to Lucas, the long phase of increasing world income inequality, discussed by Pritchett (1997), has passed. The growth rate of world production is predicted by the model to peak 'around 1970' and thereafter decline towards a rate of 2 per cent sometime just beyond the year 2100. The predictions of Lucas's model appear 'consistent with what we know about the behaviour of per capita incomes in the last two centuries' (Lucas, 2000). However, Crafts and Venables (2002) do not share the optimism of Lucas. Taking into account geographical and agglomeration factors, they conclude that the playing field is not level and therefore the convergence possibilities among the poor countries are much more limited than is suggested by Lucas. Rather, we are likely to observe the rapid convergence of a selected group of countries (for detailed and contrasting views on the evolution of global income distribution see Sala-i-Martin, 2002a, 2002b; Bourguignon and Morrisson, 2002; Milanovic, 2002).

While Solow's model predicts conditional convergence and explains growth differences in terms of 'transitional dynamics', an alternative 'catch-up' hypothesis emphasizes technological gaps between those economies behind the innovation frontier and the technologically advanced leader economies (Gerschenkron, 1962; Abramovitz, 1986, 1989, 1990, 1993). The 'catch-up'

literature also places more emphasis on historical analysis, social capability and institutional factors (see Fagerberg, 1995).

Whereas in the Solow model the main mechanism leading to differential growth rates relates to rates of capital accumulation, in the catch-up model it is the potential for low income per capita countries to adopt the technology of the more advanced countries that establishes the potential for poor countries to grow more rapidly than rich countries. In other words, there appear to be three potential (proximate) sources of growth of labour productivity, namely:

1. growth through physical and human capital accumulation;
2. growth through technological change reflecting shifts in the world production frontier;
3. growth through technological catch-up involving movement toward the world production frontier.

In other words, poor countries have the additional opportunity to grow faster by moving toward the technological frontier representing 'best practice' technology, or as P. Romer (1993) puts it, poor countries need to reduce their 'idea gaps' rather than 'object gaps'. Kumar and Russell (2002) find that there is 'substantial evidence of technological catch-up' while Parente and Prescott (2000) have emphasized that in many countries the failure to adopt 'best practice' technology is due to barriers that have been erected to protect specific groups who will be adversely affected (at least in the short run) by the changes that would result from technological change. Both the neoclassical and catch-up arguments imply that economic growth rates are likely to be closely related to per capita GDP, with poor economies benefiting in terms of economic growth from their relative backwardness. There is also accumulating evidence that more open economies converge faster than closed economies (Sachs and Warner, 1995; Krueger, 1997, 1998; Edwards, 1993, 1998; Parente and Prescott, 2000). While this appears to be true in the modern era, Baldwin et al. (2001) argue that during the Industrial Revolution international trade initially contributed to the divergence between rich and poor countries. However, they also suggest that in the modern era, the huge reduction in the transaction costs of trading ideas 'can be the key to southern industrialisation' (see also Galor and Mountford, 2003).

Finally, we should note that while there has been 'Divergence, Big Time' with respect to per capita GDP, this is in 'stark contrast' to what has been happening across the globe with respect to life expectancy, where there has been considerable convergence. Becker et al. (2003) compute a 'full income' measure for 49 developed and developing countries for the period 1965–95 that includes estimates of the monetized gains from increased longevity. By estimating economic welfare in terms of the quantity of life, as well as the

quality of life, Becker et al. show that the absence of income convergence is reversed. 'Countries starting with lower income grew more in terms of this "full income" measure. Growth rates of "full income" for the period average 140% for developed countries, and 192% for developing countries' (see also Crafts, 2003).

11.13 Beyond the Solow Model

Although the lack of a theory of technological change is a clear weakness of the basic neoclassical growth model, Mankiw (1995) argues that many general predictions from the theory are 'broadly consistent with experience'. For example, cross-country data indicate a strong negative correlation between population growth and income per capita and a strong positive correlation between income per capita and savings/investment rates (Jones, 2001a). As predicted by the model, rates of growth in the rich OECD economies are relatively low while rapid growth rates have been observed in countries moving from an initial position of relatively low income per capita and low capital intensity. There is also strong evidence of convergence among relatively homogeneous economies such as the OECD and between regions and states within the USA, Europe and Japan (Baumol, 1986; Barro and Sala-i-Martin, 1995). In larger, more diverse data sets there is little evidence of the expected negative relationship between growth rates and some initial (for example 1960) level of income per capita, that is, absolute convergence (P. Romer, 1986, 1989; DeLong, 1988). However, 'the central idea of conditional convergence receives strong support from the data' (Barro, 1991, 1997) and has considerable explanatory power for both countries and regions. The growth accounting research of Alwyn Young (1992, 1994, 1995) has shown that the rapid growth of the Asian Tiger economies is easily explicable and can be attributed mainly to rapid accumulation of factor inputs rather than unusually high total factor productivity growth. As Paul Krugman (1994b) argues, an implication of this research is that this rapid growth can therefore be expected to slow down considerably in the future, as it has already done in Japan. The Solow model has also been used to provide a plausible 'reconstruction' account of the 'miracles' of Japanese and German post-1945 growth, and also the relatively good growth performance of France and Italy, in terms of the transitional dynamics towards a high income per capita steady state. It seems plausible that these economies grew rapidly in the post-war period because they were 'reconstructing' their capital stock following the destruction resulting from the Second World War.

However, there are a number of important deficiencies and puzzles which the Solow model finds difficult to overcome and explain. First, in the Solow model, while economic policy can permanently influence the level of per

capita output (for example by raising the savings ratio via tax inducements), it cannot alter the path of long-run growth. Growth rates can only be increased temporarily during the transitional dynamics en route to the new steady state. Cross-country growth differentials are also explained in terms of the transitional dynamics which allow countries to grow faster than their long-run sustainable growth rates. Sustained growth in the Solow model is only possible if there is technological progress, since without it per capita income growth will eventually cease due to the impact of diminishing returns to capital accumulation. Given that per capita incomes have been rising for over 100 years in a large number of countries, and growth rates have displayed no overall tendency to decline, the role of technological progress in the Solow model in explaining sustainable growth becomes crucial. But herein lies the obvious shortcoming of the neoclassical model since 'the long-run per capita growth rate is determined entirely by an element – the rate of technological progress – that is outside the model … Thus we end up with a model of growth that explains everything but long-run growth, an obviously unsatisfactory situation' (Barro and Sala-i-Martin, 1995). Furthermore, as P. Romer (1989) highlights, in terms of policy advice for long-term growth the neoclassical model has little to offer!

A second problem relates to the evidence, which clearly shows that income per capita differentials across the world are much greater than predicted by the model. Differences across countries in capital intensities are too small to account for the observed disparities in real incomes. Using a Cobb–Douglas production function framework it is possible to allocate differences in the level of per capita incomes between countries to variations in levels of total factor productivity growth and the accumulation of factor inputs. In particular it is possible to estimate how much of the income disparities witnessed between rich and poor countries can be attributed to different capital intensities since total factor productivity is common across all countries. Substituting from equation (11.34) to equation (11.26) gives equation (11.38):

$$\dot{k} = sk^{\alpha} - (n + \delta)k \tag{11.38}$$

Setting this equation equal to zero (the steady state condition) and substituting into the production function yields (11.39):

$$y^{*} = [s/(n + \delta)]^{\alpha/(1-\alpha)} \tag{11.39}$$

Equation (11.39) is now in a form that enables a solution to be found for the steady state output per worker (y^{*}). As Jones (2001a) highlights, we can see from equation (11.39) why some countries are so rich and some are so poor. Assuming exogenous technology and a similar value for the capital exponent

(α), countries that sustain high rates of saving, and low rates of population growth and depreciation, will be rich. According to the neoclassical growth model the high-income economies have achieved their high living standards because they have accumulated large per worker stocks of capital. However, although the model correctly predicts the directions of the effects of saving and population growth on output per worker, it does not correctly predict the magnitudes. As Mankiw et al. (1992) and Mankiw (1995) argue, the gaps in output per worker (living standards) between rich and poor countries are much larger than plausible estimates of savings rates and population growth predict using equation (11.39). The crux of the problem is that with $\alpha = 1/3$ there are sharply diminishing returns to capital. This implies that a tenfold gap in output per worker between the USA and India would require a thousandfold difference in the capital–labour ratios between these countries! (It should be noted that this result is highly sensitive to the choice of $\alpha = 1/3$ for the share of capital in GDP.)

A third problem with the Solow model is that given a common production function (that is, exogenous technology) the marginal product of capital should be much higher in poor countries than in rich countries. Given the parameters of the Solow model, the observed tenfold differential in output per worker between rich and poor countries implies a hundredfold difference in the marginal product of capital if output gaps are entirely due to variations in capital intensities. Such differentials in the rate of return to capital are simply not observed between rich and poor countries. As David Romer (1996) observes, such differences in rates of return 'would swamp such considerations as capital market imperfections, government tax policies, fear of expropriation and so on and we would observe immense flows of capital from rich to poor countries. We do not see such flows.' But the rate of return to capital in poor countries is less than expected and the anticipated massive flows of capital from rich to poor countries have not been observed across poor countries as a whole (Lucas, 1990b).

A fourth difficulty relates to the rate of convergence, which is only about half that predicted by the model. The economy's initial conditions influence the outcome for much longer than the model says it should (Mankiw, 1995).

In conclusion, it appears that within the Solow growth framework, physical capital accumulation alone cannot account for either continuous growth of per capita income over long periods of time or the enormous geographical disparities in living standards that we observe. In terms of Figure 11.3, the data on output per worker (or income per capita) that we actually observe across the world reveal much greater disparities than those predicted by the Solow model based on differences in capital per worker.

The new growth models emerging after 1986 depart from the Solow model in three main ways. One group of models generates continuous growth by

abandoning the assumption of diminishing returns to capital accumulation. To achieve this, Paul Romer (1986) introduced positive externalities from capital accumulation so that the creation of economy-wide knowledge emerges as a by-product of the investment activity of individual firms, a case of 'learning by investing' (Barro and Sala-i-Martin, 2003). A second approach models the accumulation of knowledge as the outcome of purposeful acts by entrepreneurs seeking to maximize private profits; that is, technological progress is endogenized (P. Romer, 1990). A third class of model claims that the role of capital is much more important than is suggested by the α term in the conventional Cobb–Douglas production function shown in equations (11.28)–(11.30). In their 'augmented' Solow model, Mankiw et al. (1992) broaden the concept of capital to include 'human capital'. The first two classes of model constitute the core of endogenous growth theory whereas the Mankiw, Romer and Weil (MRW) model constitutes what Klenow and Rodriguez-Clare (1997a, 1997b) call a 'neoclassical revival'. The central proposition of endogenous growth theory is that broad capital accumulation (physical and human capital) does not experience diminishing returns. The growth process is driven by the accumulation of broad capital together with the production of new knowledge created through research and development.

11.14 Endogenous Growth: Constant Returns to Capital Accumulation

During the mid-1980s several economists, most notably Paul Romer (1986, 1987b) and Robert Lucas (1988), sought to construct alternative models of growth where the long-run growth of income per capita depends on 'investment' decisions rather than unexplained technological progress. However, as Crafts (1996) notes, the term investment in the context of these new models refers to a broader concept than the physical capital accumulation reported in the national accounts; research and development (R&D) expenditures and human capital formation may also be included. 'The key to endogenous steady state growth is that there should be constant returns to broad capital accumulation'. Hence in order to construct a simple theory of endogenous growth, the long-run tendency for capital to run into diminishing returns needs to be modified to account for the extraordinary and continuous increases in observed per capita incomes across the world's economies. In the early versions of the new endogenous growth theory the accumulation of capital plays a much greater role in the growth process than in the traditional neoclassical model. In many ways the work of Romer revives the earlier seminal contribution of Arrow (1962) on 'learning by doing'. Arrow had shown how the productivity of labour increases with experience, and experience is a function of cumulative investment expenditures that alter the work environment. That is, a firm's

accumulation of capital produces external effects on learning. However, as Blaug (2002) argues, 'it strains credulity to believe that this could account, not just for a once-and-for-all increase in output, but also for a *constant* rate of increase in total factor productivity year in year out'.

Building on Arrow's insight, Romer broadened the concept of capital to include investment in knowledge as well as the accumulation of physical capital goods. Since the knowledge gained by workers in one firm has public good characteristics and is at best only partially excludable, then knowledge spillovers occur such that investment in knowledge (R&D) by one firm increases the production potential of other firms. No individual firm can completely internalize the positive impact that their investment in physical and human capital has on the economy-wide stock of knowledge.

Paul Romer's 1986 model can be illustrated by modifying the production function. In equation (11.40) the production function includes technology (A) as an endogenous input:

$$Y = F(K, L, A) \tag{11.40}$$

At the micro level, the output of an individual firm (j) depends on its own inputs of capital (Kj), labour (Lj) and the economy-wide state of knowledge (A), as indicated in equation (11.41):

$$Yj = F(Kj, Lj, A) \tag{11.41}$$

In this formulation the growth of knowledge (technology) is assumed to depend on the growth of capital because capital deepening fosters technological spillovers that raise the marginal productivity of capital across the economy as a whole. Therefore any increase in aggregate K will improve A and hence the productivity of all firms. In Romer's (1986) endogenous growth model the expansion of aggregate knowledge results from learning externalities among firms. In effect, the higher the level of the capital stock in an economy, the more productive each firm will be via a process of 'learning by doing'. So while a firm's production function exhibits constant returns to scale and diminishing returns to capital accumulation, the aggregate production function will exhibit increasing, rather than constant, returns to scale.

One of the simplest models of endogenous growth is the AK^* model shown in equation (11.42) below (Rebelo, 1991):

$$Y = K^{\alpha} H^{\beta} = AK^* \tag{11.42}$$

Here A is a constant, K^* represents a broad measure of capital ($K^{\alpha} H^{\beta}$), and $\alpha + \beta = 1$. As Crafts (1995) points out, 'models of this kind put investment

centre stage and see growth as an investment-driven process. There is no role for the Solow residual.' Therefore there is a close similarity between the *AK* model and the Harrod–Domar model. In both models there are no diminishing returns and hence no reason for growth to slow down as capital deepening occurs. If one group of countries has higher average savings rates, lower depreciation rates and lower capital–output ratios than some other group of countries, then the first group will grow faster than the second group permanently and 'divergence, big time' will be the rule.

The *AK* class of endogenous growth models has been subject to heavy criticism, mainly on account of their key assumption of an absence of diminishing returns to the capital input. The *AK* model predicts a permanent increase in the growth rate following an increase in the investment/GDP ratio of an economy. However, Jones (1995), in a time series analysis of 15 OECD countries in the post-1945 period, argues that the *AK* models are inconsistent with the empirical evidence. Although the investment/GDP ratios increased significantly in the 1950–89 period, growth rates of GDP per worker remained stable or have fallen. This finding has been challenged by McGrattan (1998). By considering time series evidence from a larger sample of countries over a longer time period McGrattan finds the main predictions of *AK* theory to be confirmed by the data. Using data from Maddison (1995) for the period 1870–1989, McGratton finds that 'higher investment rates correspond to higher growth rates, with the exception of the US economy where there is little variation in the growth rate of GDP per capita'. Extending the analysis to cross-sectional data for 125 economies in the period 1960–85 also reveals 'a definite positive correlation between investment rates and growth rates'.

11.15 Endogenous Growth: The Economics of Ideas

The issue of convergence has raised as an important question the possible importance of differences in technology (knowledge) across the economies of the world. The Solow model attempts to explain per capita income level and growth differences assuming that technology is a pure public good and is therefore freely available to all countries irrespective of their level of development. An increasing number of economists, and most economic historians and development economists see significant technology gaps as the crucial problem facing poor countries. Such an approach emphasizes the need for policies to be adopted that will close 'idea gaps' between nations (P. Romer, 1993).

Paul Romer's 1986 model explains technological progress as an unintentional by-product of capital accumulation by individual firms. Subsequently, Romer (1990), dissatisfied with his initial approach, proceeded to develop a second strand of new growth theory. Endogenous innovation models embrace

a neo-Schumpeterian framework of endogenous technological change based on three premises (Grossman and Helpman, 1991, 1994; Crafts, 1996; Aghion and Howitt, 1998). First, as in the Solow model, the basic driving force behind economic growth is technological change, that is, improvements in knowledge about how we transform inputs into outputs in the production process. Second, technological change is endogenous, being determined by the deliberate activities of economic agents acting largely in response to financial incentives. Third, the defining characteristic of ideas/knowledge is that 'once the cost of creating a new set of instructions has been incurred, the instructions can be used over and over again at no additional cost' (Romer, 1990). Therefore ideas are non-rivalrous outputs and their use by one firm or person does not in any way reduce their availability to other firms or persons. Ideas are also 'partially excludable', where excludability is defined as the ability of the owners of a good to prevent other economic agents from using it without payment. As Romer (1990) notes, 'excludability is a function of the technology and the legal system'. Given Romer's second premise that technological change results from the purposeful actions of self-interested economic agents, improvements in technology (new ideas) must generate benefits to individuals that are at least 'partially excludable', for example by having patent laws.

Romer's insights have led to a burgeoning of research into the economics of ideas (Jones, 2002, 2005). The three premises discussed above have two important implications for the theory of economic growth. First, because ideas are non-rivalrous, they can be accumulated without limit on a per capita basis. Second, because of incomplete excludability (appropriability), knowledge creation involves substantial spillovers of benefits (externalities) which cannot be entirely captured by the economic agents who produce the ideas. The 'unbounded' growth and 'incomplete appropriability' features of the economics of ideas imply that 'output cannot be a constant-returns-to-scale function of all its inputs taken together'. Romer's analysis implies increasing returns to scale and by implication microfoundations based on the presence of imperfect competition (see Romer, 1994a).

While a non-rivalrous good such as a new idea involves a fixed cost of production, which is often substantial, once the new knowledge has been created there is zero marginal cost involved with any further use of the new idea. A new design is costly to produce but once in existence it can be used as often as desired and in as many contexts as desired. It is for this reason that legal mechanisms such as patents and copyrights exist in order to grant investors monopoly rights over a new idea, at least for a time, so that they can earn a reward for their new ideas (Kremer, 1998; Mazzoleni and Nelson, 1998). The importance of this issue has been illustrated by North (1990), who argues that the economic development of Western Europe did not seriously

begin until the development of property rights ensured that individuals could reap some of the benefits of their 'ideas' and helped to speed up the pace of technological change (Crafts, 1995). The era of modern economic growth, beginning with the Industrial Revolution in Britain,

> occurred when the institutions protecting intellectual property rights were sufficiently well developed that entrepreneurs could capture as a private return some of the enormous social returns their innovations would create … history suggests that it is only when the market incentives were sufficient that widespread innovation and growth took hold. (Jones, 2001a)

In the case of the USA the framers of the US Constitution were eager to 'promote the progress of science and useful arts'. Therefore an intellectual property clause providing for copyright and patent rights appears in the first article of the Constitution and by 1810 the USA 'far surpassed Britain in patenting per capita' (Khan and Sokoloff, 2001). The failure of China to lead the first Industrial Revolution has also been attributed to that country's inability to establish a free market, institutionalize property rights, provide an environment conducive to emulation and innovation, and to absorb foreign technology (Landes, 1998). Thus according to the new breed of endogenous growth models, 'the government has great potential for good or ill through its influence on the long-term rate of growth' (Barro, 1997). Economic growth can be influenced not only by policies that affect trade regimes, technology transfer, the provision of infrastructure and financial markets, but also by policies that affect taxation and incentives, the protection of intellectual property rights and the maintenance of law and order.

By developing an endogenous theory of technological change Romer has challenged both the traditional and augmented versions of the Solow neoclassical growth model (see below, section 11.16). In the neoclassical model technology is assumed to be exogenous and hence available without limitation everywhere across the globe. Romer (1995) rejects this assumption on the basis of 'overwhelming evidence' that technology is not a pure public good. The neoclassical model emphasizes 'object gaps', differences in physical and human capital, in explaining income per capita differentials across nations. While Mankiw (1995) believes that much of the variation in living standards can be explained by differences in the quantities of human and physical capital, in contrast Romer (1993) emphasizes 'idea gaps', productivity differences resulting from technology gaps, as the main source of divergent living standards.

Parente and Prescott (1994, 1999, 2000) also attribute differences in international incomes to technology gaps. In their research they have found evidence to suggest that these productivity gaps are not caused by fundamental differences in the stock of available knowledge that developing countries have

access to. Instead, Parente and Prescott argue that there exist barriers in the form of society-imposed constraints which prevent firms in many developing countries from adopting better production methods, and many of these constraints 'are put in place to protect the interests of groups vested in current production processes'. As a result they conclude that most differences in international incomes 'are the result of differences in total factor productivity'. Parente and Prescott (2005) conclude that 'changes in a country's institutions that result in large increases in the efficiency with which resources can be used in production give rise to growth miracles'.

Romer's position has received recent support from the research of Easterly and Levine (2001), who find that the 'residual' (total factor productivity) rather than factor accumulation can explain most of the cross-country income and growth differentials. Their data show that while factor accumulation is persistent, growth is not. Nelson and Pack (1999), in their discussion of the Asian miracle and modern growth theory, also stress the importance of the entrepreneurship, innovation and learning that these economies had to undertake before they could successfully assimilate new technologies. In their view the accumulation of human and physical capital is a necessary but far from sufficient part of this process. What is crucial for success is the establishment of a policy environment that nurtures learning, and for economists to better understand the learning process taking place during the assimilation of new ideas and technologies they need 'a better theory of firm behaviour in such situations'.

Historical experience demonstrates that the creation and transmission of ideas has undoubtedly been an important determinant of current living standards (Rosenberg, 1994; Mokyr, 2005). If Romer is correct and the poor countries do suffer from idea gaps rather than object gaps, then a significant part of worldwide poverty can be eliminated 'at relatively low cost' via technological 'catch-up'. A clear implication of this analysis is that nations which isolate themselves from the free flow of ideas, or erect barriers to the adoption of new technologies, will suffer relative stagnation since trade policies and openness affect innovation and growth. Foreign direct investment can act as a significant channel for the diffusion of new innovations and ideas, thereby enhancing the growth process (Grossman and Helpman, 1990; Romer, 1994b; Sachs and Warner, 1995; Proudman and Redding, 1997; Edwards, 1998; Parente and Prescott, 2000). Therefore, at least potentially, poor economies have the most to gain from reducing restrictions to international trade, encouraging inward FDI flows and investing in human capital because by doing so they can gain access to the stock of world knowledge (World Bank, 1998/9). While in the neoclassical model the removal of inefficiencies caused by trade barriers will produce level effects on production possibilities but no sustained growth effects, in endogenous growth models

the growth effects of increasing economic integration are likely to be much more important.

A further implication of Paul Romer's research is that for the USA to maintain its leadership position, government policies must continue to support a high level of R&D activities in both private and public institutions. Given the well-documented large divergence between social and private rates of return from R&D expenditures, the government has a vital role to play in preventing underinvestment in this activity. In a recent investigation of the optimal rate of R&D investment in the USA Jones and Williams (1998) conclude that the private rate of return to R&D in the USA is of the order of 7–14 per cent, while a 'conservative estimate' of the social rate of return is 30 per cent. Therefore Jones and Williams conclude that optimal R&D spending as a share of GDP is 'more than two to four times larger than actual spending' (see also Jones and Williams, 2000).

In contrast to the supply-side view of the growth of knowledge, ideas and technological change, Schmookler (1966) argues that technological change is primarily demand-induced. Unlike Romer's model, where a key input to the development of new technology is the supply of previous innovations (see Jones, 2005), Schmookler sees the stimulus to technological change and innovation as the need to solve current technological problems; that is, technological change is demand-driven and dependent on the usefulness of new ideas. In a recent discussion of Schmookler's work, Kelly (2002) concludes that the supply- and demand-side influences on technology are complementary.

An important deficiency of recent endogenous growth theories is that they lose the prediction of conditional convergence, a prediction which Barro (1997) argues has a 'strong empirical regularity in the data for countries and regions'. To rectify this flaw Barro and Sala-i-Martin (1997) have developed a model that combines elements of endogenous growth with the convergence implications of the Solow model. Their model has the following elements:

1. in the long run the rate of growth in the world economy is driven by technological discoveries in the leading economies;
2. follower economies share in the new innovations via a process of imitation;
3. since imitation is generally cheaper than innovation, 'most countries prefer to copy rather than invent';
4. the relatively low cost of imitation implies that the follower economies will grow relatively faster than the leader economies and converge, at least part way, towards the leaders;
5. as the amount of uncopied innovations decreases, the costs of imitation will tend to rise and therefore the follower's growth rate will tend to slow down;

6. therefore, the Barro/Sala-i-Martin model generates a form of conditional convergence based on the diffusion of technology across countries and resembles the predictions of the Solow model;
7. in the long run 'all economies grow at the rate of discovery in the leading places'.

The Barro/Sala-i-Martin hybrid model therefore establishes a framework where long-run growth is driven endogenously by the discovery of new ideas in the 'leading-edge' economies, but also retains the convergence properties of the neoclassical growth model via the impact of the imitation behaviour of follower countries.

11.16 An Augmented Solow Model: A Neoclassical Revival?

As it stands, the neoclassical growth model, relying as it does on differences in capital–labour ratios across countries to explain the wide disparities in levels of per capita output, cannot satisfactorily explain world income differentials. In response to this deficiency Mankiw et al. (1992) 'augment' the Solow model by including the accumulation of human capital as well as physical capital. The key to their approach is the argument that the conventional estimate of α, capital's income share, may not be a good indicator of the overall contribution of capital. By adding human capital to the model the production function becomes (11.43):

$$Y = K^{\alpha} H^{\beta} (AL)^{1-\alpha-\beta} \text{ and } \alpha + \beta < 1 \tag{11.43}$$

Here we now have four factors of production combining to produce output where H is the stock of human capital and AL is the labour input measured in efficiency units, which captures both the quantity of labour and the productivity of labour determined by available technology (see Mankiw, 2003). The production function exhibits constant returns to scale and with $\alpha + \beta < 1$ there are diminishing returns to 'broad capital'. But with a larger capital share ($\alpha + \beta = 2/3$) the average product of labour declines more slowly as accumulation takes place since the size of the capital share determines the curvature of the production function and hence the speed at which diminishing returns set in. Diminishing returns to the broader concept of capital will be much less severe than in the traditional Solow model where $\alpha = 1/3$. When α is small, the curvature of the production function in Figure 11.3 is large. But by augmenting the model with human capital, the transition to the steady state is much slower and 80 per cent of international differences in living standards can be explained by differences in the rate of population growth and the accumulation of both human and physical capital (Mankiw et al., 1992;

Mankiw, 1995). The transitory impact of any increase in the rate of investment in the MRW model will have prolonged effects. However, because the exponents on *K* and *H* sum to less than one, this 'neoclassical revival' in growth theory does not provide a model of endogenous growth. Per capita income will eventually settle down in a steady state and grow at the exogenously determined rate of technological progress.

For some critics the MRW model, by taking the public-good view of technology, has failed to address the crucial issue of variations in total factor productivity growth and technical efficiency across nations (Klenow amd Rodriguez-Clare, 1997a; 1997b). While the augmented Solow model better explains international differences in living standards, it cannot account for the persistence of economic growth. Endogenous growth theory attempts to show how persistent growth may take place without having to resort to exogenous technological progress (Bernanke and Gurkaynak, 2001).

11.17 Focusing on the Fundamental Causes of Growth

The research of economists shows that successful economies are those with high rates of accumulation of human and physical capital together with sustained technological progress. But this conclusion then raises the crucial question: why do some nations successfully achieve this outcome while others fail? Olson (1996) has highlighted the fact that high rates of growth seem to occur in a subset of poor countries rather than in all low-income countries as the transitional dynamics of the Solow neoclassical growth model imply. Given that capital and technology can migrate across political boundaries, the persistence of significant differences in the level of output per worker suggests the presence of persistent barriers to growth and development (Parente and Prescott, 2000). An obvious deterrent to the free flow of capital from rich to poor countries arises from the greater risk involved in investing in countries characterized by macroeconomic instability, trade barriers, inadequate infrastructure, poor education, ethnic diversity, widespread corruption, political instability, disadvantageous geography and frequent policy reversals. To understand why some countries have performed so much better than others with respect to growth it is therefore necessary to go beyond the proximate causes of growth and delve into the wider fundamental determinants. This implies that we cannot hope to find the key determinants of economic growth by using narrow economic analysis alone. To explain growth 'miracles' and 'disasters' requires an understanding of the history of the countries being investigated as well as how policy choices are made within an institutional structure involving political distortions.

Dani Rodrik (2003) has provided a useful framework for highlighting the distinction between the proximate and fundamental determinants of eco-

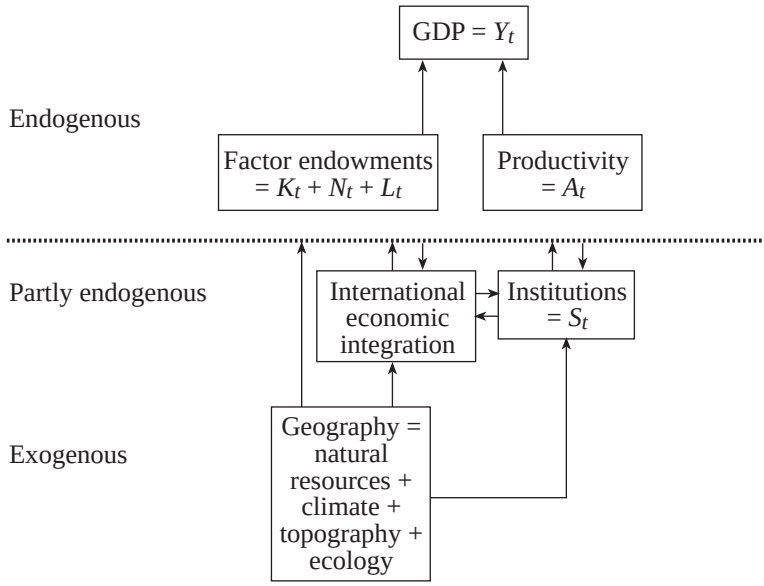


Figure 11.8 Proximate and fundamental sources of growth

nomic growth. Figure 11.8, adapted from Rodrik, captures the main factors that determine the size and growth of any economy. Referring back to equation (11.4), in the upper part of Figure 11.8 we can see the influence of the proximate determinants of growth, with output being directly influenced by an economy's endowments of labour (L_t), physical capital (K_t), natural resources (N_t) and the productivity of these resources (A_t). The impact of both technical and allocative efficiency is captured within the productivity variable. In the lower portion of Figure 11.8 we observe the major fundamental determinants of economic growth, including social capability (S_t). Rodrik provides a threefold taxonomy of the fundamental determinants of growth, namely geography, integration and institutions. These categories highlight three major research areas, within a voluminous and rapidly expanding literature, that have dominated growth analysis in recent years. Many social scientists would argue forcefully that the influence of culture should be added to the list of important deeper determinants of economic performance. It is certainly the case that economic historians have given much greater consideration to culture as a determinant of economic performance than economists. For example, David Landes argues that 'Culture Makes Almost All the Difference' (Harrison and Huntington, 2000). For other interesting discussions of the influence of culture on economic growth and development the reader should consult

Huntington, 1996; Temin, 1997; Landes, 1998; Lal, 1999; Dasgupta and Serageldin, 2000; Barro and McCleary, 2003; Grief, 2003.

As Rodrik points out, the central question in growth analysis is: which of the causal relationships in Figure 11.8 matters most? However, Rodrik also notes that geography is the only exogenous factor in his threefold taxonomy, with integration and institutions 'co-evolving with economic performance'. The causal interrelationships between the variables in Figure 11.8, indicated by the two-way direction of some of the arrows, suggest that there are complex feedback effects at work. Therefore empirical work, in the form of endless cross-country regressions, that attempts to establish clear lines of causality must be treated with 'extreme care'.

11.18 Institutions and Economic Growth

While poor countries have enormous potential for catch-up and convergence, these advantages will fail to generate positive results on growth in countries with an inadequate political, legal and regulatory framework. The notion that institutions profoundly influence the wealth of nations is of course an old idea first eloquently expressed by Adam Smith. Ever since the publication of Adam Smith's *Wealth of Nations* in 1776 economists have been aware that security of property rights against expropriation by fellow citizens or the state is an important condition for encouraging individuals to invest and accumulate capital. Given this pedigree, economists have tended to centre their analysis of the deeper determinants of growth on the role of institutions. Emphasis here is placed on factors such as the role of property rights, the effectiveness of the legal system, corruption, regulatory structures and the quality of governance (North, 1990; World Bank, 1997; Olson, 2000; Acemoglu et al., 2001, 2002a; Glaeser and Shleifer, 2002). From an institutionalist perspective the search for the deeper determinants of growth has led to what Hibbs (2001) has called 'the politicisation of growth theory'. By 'politicisation' Hibbs is referring to the increasing emphasis given by many growth researchers, in particular the seminal contributions of Douglass North, to the importance of 'politics, policy and institutional arrangements'. These factors ultimately determine the structure of incentives, the ability and willingness of people to save and invest productively, the security of property rights and the incentive to innovate and participate in entrepreneurial activity. Political–institutional factors also appear to be robust determinants of growth in many cross-country regression studies (Knack and Keefer, 1995, 1997a, 1997b; Sala-i-Martin, 1997; Dawson, 1998; Durlauf and Quah, 1999; Easterly and Levine, 2003; Rodrik, 2003; Rodrik et al., 2004).

While the presence of technological backwardness and income per capita gaps creates the potential for catch-up and convergence, Abramovitz (1986) has highlighted the importance of 'social capability' without which countries

will not be able to realize their potential. Social capability refers to the various institutional arrangements which set the framework for the conduct of productive economic activities and without which market economies cannot function efficiently. Temple and Johnson (1998) suggest that indexes of social development developed in the 1960s by Adelman and Morris (1967) have 'considerable predictive power' (see also Temple, 1998). Temple and Johnson (1998) show that these measures, after allowing for initial income, are 'very useful in predicting subsequent growth' and 'if observers in the early 1960s had given more emphasis to these indexes of social capability, they might have been rather more successful in predicting the fast growth of East Asia, and underperformance of sub-Saharan Africa'.

Developing effective institutions

Clearly the major income differentials that we observe around the world have a lot to do with differences in the quality of countries' institutions and economic policies as well as the quality of political leadership. This explains the rapid growth witnessed in a subset of East Asian developing countries since around 1960 and the relative stagnation of most of sub-Saharan Africa over that same period. Many economists believe that the main reasons why some countries progress and grow rapidly while others stagnate cannot be found in the area of geography and factor endowments. Countries with poor natural resources such as Japan, Taiwan and South Korea have experienced 'miracle' growth while many natural resource-abundant economies in sub-Saharan Africa, such as Zaire (from 1997 the Democratic Republic of Congo), have been growth 'disasters'. Reynolds (1985) concludes that the single most important explanatory variable of economic progress is the political organization and the administrative competence of government (see Herbst, 2000). Gray and McPherson (2001), in their analysis of the leadership factor in African policy reform, conclude that 'a large number of African countries, perhaps the majority, have been ruled by individuals who had sufficient power to implement reforms had they been so motivated. However, their motivation led them in different directions.' It is also the case that the political and economic losers from the changes that economic development requires will often act as barriers to progress (Acemoglu and Robinson, 2000a; Parente and Prescott, 2000).

Given the importance of these issues, the recent political economy of growth literature has focused on such factors as the relationship between economic freedom, democracy and growth (for example Barro, 1996, 1999; Clague et al., 1996; Minier, 1998; Benson Durham, 1999; Landman, 1999; Olson, 2000); property rights and growth (for example North and Weingast, 1989; North, 1990; DeLong and Shleifer, 1993; Acemoglu and Johnson, 2003); ethnic heterogeneity, political conflict and growth (for example East-

erly and Levine, 1997; Acemoglu and Robinson, 2000a, 2000b, 2000c, 2001, 2003; Collier, 2001; Easterly, 2001b); the impact of inequality and political instability on growth (for example Alesina and Rodrik, 1994; Alesina et al., 1996; Lee and Roemer, 1998; Barro, 2000; Glaeser et al., 2003; see also Chapter 10); and various measures of social capability/social infrastructure/social capital and growth, including trust (for example Knack and Keefer, 1995, 1997a, 1997b; Abramovitz and David, 1996; Landau et al., 1996; Hall and Jones, 1997, 1999; Temple, 1998; Temple and Johnson, 1998; Paldam, 2000; Rose-Ackerman, 2001; Zak and Knack, 2001).

According to the World Bank (2002), there is a growing body of evidence linking the quality of institutional development to economic growth and efficiency across both time and space and there is now widespread acceptance of the idea that 'good' institutions and incentive structures are an important precondition for successful growth and development. Because economic history is essentially about the performance of economies over long periods of time, it has a significant contribution to make in helping growth theorists improve their ability to develop a better analytical framework for understanding long-run economic change (North and Thomas, 1973; North, 1981, 1989, 1990, 1994; Myles, 2000).

The story that emerges from economic history is one which shows that the unsuccessful economies, in terms of achieving sustained growth of living standards, are those that fail to produce a set of enforceable economic rules of the game that promote economic progress. As North (1991) argues, the 'central issue of economic history and of economic development is to account for the evolution of political and economic institutions that create an economic environment that induces increasing productivity'.

North (1991) defines institutions as 'the humanly devised constraints that structure political, economic and social interaction'. The constraining institutions may be informal (customs, traditions, taboos, conventions, self-imposed codes of conduct involving guilt and shame) and/or formal (laws, contract enforcement, rules, constitutions, property rights). In an ideal world the informal and formal institutions will complement each other. These institutions provide a structure within which repeated human interaction can take place, they support market transactions, they help to transmit information between economic agents and they give people the incentives necessary to engage in productive activities. History is 'largely a story of institutional evolution' and effective institutions 'raise the benefit of co-operative solutions or the costs of defection' (North, 1991).

A good example demonstrating the importance of institutions for sustained economic growth is provided by the post-Second World War reconstruction of Europe. As DeLong and Eichengreen (1993) argue, 'the Marshall Plan significantly sped western European growth by altering the environment in

which economic policy was made' and by providing support to a recovery strategy based on the restoration of a market-based economic system, together with the necessary supporting institutions. In retrospect we now know that the period 1950–73 turned out to be a 'Golden Age' of economic growth in the 'mixed' economies of Western Europe, and DeLong and Eichengreen conclude that the Marshall plan was 'history's most successful structural adjustment programme'. Eichengreen (1996) also extends the institutional based explanation of why Europe was able to enjoy a 'Golden Age' of economic growth in the 25-year period following the implementation of the Marshall Plan. European economic growth during this quarter-century was faster than any period either before or since (Maddison, 2001). According to Eichengreen, the foundation for this 'Golden Age' was a set of domestic (the social market economy) and international institutions (GATT, the development of free intra-European trade, the Bretton Woods institutions) that 'solved problems of commitment and co-operation that would have otherwise hindered the resumption of growth'.

For individuals living in a typical rich OECD economy in the twenty-first century it is easy to take most of these market-based institutions for granted because they have evolved over such a long historical period. But the 'trials of transition' witnessed in the former communist economies remind us just how difficult it is to make market economies operate effectively without having the necessary institutional infrastructure in place.

Evidence from 'natural experiments'

One very important source of divergence in per capita incomes emphasized by Fukuyama (1989, 1992), Olson (1996, 2000) and DeLong (2001) has arisen because of political developments which have influenced the choice of economic system and policies. Those countries which attempted to 'develop' behind the 'Iron Curtain' now have much lower income per capita than countries which had a comparable income per capita in 1950 and followed the capitalist path.

> The fact that a large part of the globe was under communist rule in the twentieth century is one major reason for the world's divergence … depending on how you count and how unlucky you are, 40 to 94 per cent of the potential material prosperity of a country was annihilated if it happened to fall under communist rule in the twentieth century. (DeLong, 2001)

The most obvious examples involve the comparative development experiences of East and West Germany, North and South Korea, and China with Taiwan/Singapore/Hong Kong. But comparisons between other neighbouring countries seems reasonable, for example, comparisons between Russia and Finland, Hungary and Austria, Greece and Bulgaria, Slovenia and Italy, and Cambodia and Thailand reveal significant differences in living standards.

Of the examples mentioned above, the most dramatic 'natural experiment' has occurred in the Korean peninsula during the second half of the twentieth century. Following the surrender of Japan in August 1945, Korea was divided at the 38th parallel into two zones of occupation, with armed forces from the Soviet Union occupying the 'North' and American armed forces occupying the 'South'. In the summer of 1948, following the May elections, the American zone of occupation became the Republic of Korea, and in September 1948 the northern zone became formally known as the Democratic People's Republic of North Korea. Both 'Koreas' claimed full political jurisdiction over the entire Korean peninsula and this disagreement led to the Korean War, which lasted from June 1950 until the armistice of July 1953. Since then the 38th parallel has remained the dividing line between the two Koreas, with the 'communist North' adopting a centrally planned economic strategy and the 'capitalist South' putting its faith in a capitalist mixed economy. As the data in Tables 11.4 and 11.5 make clear, the impact of these choices on living standards in the two Koreas, made some 50 years ago, could not have been more dramatic. As Acemoglu (2003b) notes, 'a distinguishing feature of Korea before separation was its ethnic, linguistic and economic homogeneity. The north and south are inhabited by essentially the same people with the

Table 11.4 A tale of two Koreas

Indicator	*Population ('000)*	*GDP PPP $ millions*	*GDP per capita PPP $*	*Population ('000)*	*GDP PPP $ millions*	*GDP per capita PPP $*
Year	*North Korea*	*North Korea*	*North Korea*	*South Korea*	*South Korea*	*South Korea*
1950	9 471	7 293	770	20 846	16 045	770
1955	8 839	9 361	1 054	21 552	22 708	1 054
1960	10 392	11 483	1 105	24 784	27 398	1 105
1965	11 869	15 370	1 295	28 705	37 166	1 295
1970	13 912	27 184	1 954	32 241	62 988	1 954
1975	15 801	44 891	2 841	35 281	111 548	3 162
1980	17 114	48 621	2 841	38 124	156 846	4 114
1985	18 481	52 505	2 841	40 806	231 386	5 670
1990	20 019	56 874	2 841	42 869	373 150	8 704
1995	21 553	32 758	1 520	45 081	534 517	11 873
1998	21 234	25 131	1 183	46 430	564 211	12 152

Source: Adapted from Maddison (2001).

Table 11.5 Growth rates of per capita GDP (%): the two Koreas

	1950–73	*1973–98*
North Korea	5.84	–3.44
South Korea	5.84	5.99

Source: Adapted from Maddison (2001).

same culture, and there were only minor differences between the two areas.' Therefore, this natural experiment, of dividing the Korean peninsula into two countries, each distinguished by very different policies and institutions, 'gives a clear example of how, despite the very similar economic conditions, political leaders often chose very different policies with very different outcomes'.

As Maddison's (2001) data indicate, per capita GDP in North Korea in 1950 was $770 (at 1990 international prices). By 1998 this had only risen to $1183. In sharp contrast, although per capita income in South Korea in 1950 was also $770, by 1998 it had risen to $12 152! This again demonstrates the powerful effect that differential growth rates can have on the relative living standards of two countries. While North Korean per capita economic growth during the 1950–73 period was initially impressive (5.84 per cent per annum), during the 1973–98 period the rate of growth collapsed to minus 3.44 per cent. During the 1950–73 period South Korea's growth rate was also 5.84 per cent. However, during the years 1973–98 South Korea's per capita growth rate increased to 5.99 per cent. By 1999, World Bank (2002) data indicate that the 47 million people living in the South had a life expectancy of 73 whereas for the 23.6 million people living in the North, life expectancy was 60 and in recent years North Korea has been experiencing a famine (Noland et al., 2001).

These 'natural experiments' show that where national borders also mark the boundaries of public policies and institutions, easily observable differentials in economic performance emerge (Fukuyama, 1992; Olson, 1996). In DeLong's (1992) view, 'over the course of the twentieth century communism has been a major factor making for divergence: making nations that were relatively poor poorer even as rich industrial economies have grown richer'.

Democracy, the quality of governance and growth

Does growth promote democracy or does democracy promote growth? Recent research into the link between democracy, dictatorship and growth has produced support for both of the above linkages. Barro (1996, 1997, 1999) provides evidence in support of the Lipset (1959) hypothesis which suggests that prosperity promotes democracy. Barro's research confirms this hypoth-

esis as a 'strong empirical regularity'. Since the empirical evidence also supports the hypothesis that economic freedom promotes prosperity, Barro concludes that policies that promote economic freedom will also promote greater democracy through the Lipset prosperity effect. It is certainly indisputable that there has never been a liberal democracy (free and regular competitive elections) where there is an absence of economic freedom (see Friedman, 1962; Kornai, 2000; Snowdon 2003b).

Bhagwati's (1995) essay on democracy rejects an earlier popular view highlighted in his 1966 book on *The Economics of Underdeveloped Countries* that developing countries may face a 'cruel dilemma' in that they must somehow choose between economic development or democracy. During the 1960s democracy was often portrayed as a luxury that poor countries could not afford. It was often argued that to achieve rapid growth requires tough decisions and that in turn necessitates firm political leadership free from democratic constraints. The balance of opinion has now moved away from accepting as inevitable this 'Cruel Dilemma Thesis'.

While prosperity undoubtedly sows the seeds of democracy, the idea that a stable democracy is good for sustained growth has also been receiving increasing support in the literature. If property rights are the key to reducing transaction costs and the promotion of specialization and trade, then it should be no surprise to observe that 'almost all of the countries that have enjoyed good economic performance across generations are countries that have stable democratic governments' (Olson, 2000; Rodrik, 2000). Whereas good governance and economic prosperity are good bedfellows, autocrats, who are also invariably kleptomaniacs, are a high-risk form of investment. As Easterly (2001a) notes, 'governments can kill growth'.

For most of human history the vast majority of the peoples of the world have been governed by what Mancur Olson (1993, 2000) calls 'roving bandits' and 'stationary bandits'. History provides incontrovertible evidence that benevolent despots are a rare breed. Roving bandits (warlords) have little interest in promoting the well-being of the people living within their domain. A territory dominated by competing roving bandits represents a situation of pure anarchy and any form of sustainable economic development is impossible. With no secure property rights there is little incentive for people to produce any more than is necessary for their survival since any surplus will be expropriated by force. Stationary bandits, however, can extract more tax revenue from the territory they dominate if a stable and productive economy can be encouraged and maintained. In this situation despots have an incentive to provide key public goods such as law and order. But property rights can never be fully secure under autocratic forms of governance because the discretionary powers of the autocrat create a time-inconsistency problem. That is, the autocrat will always have a credibility problem. History shows

that absolutist princes always find it difficult to establish stable dynasties, and this uncertainty relating to succession prevents autocrats from taking a longer-term view of the economy. For example, the monarchy in England between the rule of William the Conqueror (1066) and the 'Glorious Revolution' (1688) was plagued by repeated crises of succession (for example the 'Wars of the Roses'). Only in a secure democracy, where representative government is accountable and respectful of individual rights, can we expect to observe an environment created that is conducive to lasting property rights (Fukuyama, 1989, 1992).

Acemoglu's recent research highlights the importance of 'political barriers to development'. This work focuses on attitudes to change in hierarchical societies. Economists recognize that economic growth is a necessary condition for the elimination of poverty and sustainable increases in living standards. Furthermore, technological change and innovation are key factors in promoting growth. So why do political élites deliberately block the adoption of institutions and policies that would help to eliminate economic backwardness? Acemoglu and Robinson (2000a, 2000b, 2000c, 2001, 2003, 2005) argue that superior institutions and technologies are resisted because they may reduce the political power of the elite. Moreover, the absence of a 'strong institution' allows autocratic rulers to 'adopt political strategies which are highly effective at diffusing any opposition to their regime … the kleptocratic ruler intensifies the collective action problem and destroys the coalition against him by bribing the pivotal groups' (Acemoglu et al., 2003b). Often financed by natural resource abundance and foreign aid, kleptocrats follow an effective power sustaining strategy of 'divide and rule'. In the case of Zaire, with over 200 ethnic groups, Mobutu was able to follow such a strategy from 1965 until he was overthrown in 1997. Acemoglu's research reinforces the conclusions of Easterly and Levine (1997), who find that ethnic diversity in Africa reduces the rate of economic growth (see next section).

The general thesis advocated by North and Olson is also confirmed by DeLong and Shleifer (1993), who show that those cities in medieval Europe that were under more democratic forms of government were much more productive than those under the autocratic rule of 'princes'. The incompatibility of despotism with sustainable economic development arises because of the insecurity of property rights in environments where there are no constitutional restrictions on an autocratic ruler. DeLong and Shleifer assume that the size of urban populations is a useful proxy for commercial prosperity and 'use the number and sizes of large pre-industrial cities as an index of economic activity, and changes in the number of cities and the sizes of urban population as indicators of economic growth'. Their city data show how between the years 1000 and 1500, the centre of economic gravity in Europe

moved steadily northward. Although in the year 1000 Western Europe was a 'backwater' in terms of urban development, by 1800 it was established as the most prosperous and economically advanced region of the world. While London is ranked as the 25th largest European city at the beginning of the thirteenth century, by 1650 it had risen to second place (after Paris), and by 1800 London was first. DeLong and Shleifer argue that security of property can be thought of as a form of lower taxation, with the difference between absolutist and non-absolutist governments showing up as different tax rates on private property. It has also been argued by Douglass North that the establishment of a credible and sustainable commitment to the security of property rights in England required the establishment of parliamentary supremacy over the crown. This was achieved following the 'Glorious Revolution' of 1688 which facilitated the gradual establishment of economic institutions conducive to increasing security in property rights (North and Weingast, 1989; North, 1990). The contrasting economic fortunes of the North and South American continents also bear testimony to the consequences of divergent institutional paths for political and economic performance (Sokoloff and Engerman, 2000; Acemoglu et al., 2001, 2002a; Khan and Sokoloff, 2001).

The failure in many countries to develop good governance has had serious, often drastic, economic and political consequences. The case for democracy rests very much on how regular elections and a free press and media act as important mechanisms that increase the accountability of politicians. In a principal–agent model, the scope for rent-seeking activities by politicians, who have potential access to huge resources, is greatly increased if the citizens of a country lack information and are denied the opportunity to hold politicians accountable via regular and guaranteed competitive elections (see Adsera et al., 2003).

In sum, we think that Winston Churchill had it right when he made his famous statement defending democracy in the House of Commons (11 November 1947):

> No one pretends that democracy is perfect or all wise. Indeed it has been said that democracy is the worst form of government except for all those other forms that have been tried from time to time.

Rent seeking, trust, corruption and growth

In order to foster high levels of output per worker, social institutions must be developed which protect the output of individual productive units from diversion. Countries with perverse infrastructure, such as a corrupt bureaucracy, generate rent-seeking activities devoted to the diversion of resources rather than productive activities such as capital accumulation, skills acquisition, and the development of new goods and production techniques (Murphy et al.,

1993; Mauro, 1995). In an environment of weak law and contract enforcement, poor protection of property rights, confiscatory taxation and widespread corruption, unproductive profit- (rent-) seeking activities will become endemic and cause immense damage to innovation and other growth-enhancing activities (Tanzi, 1998).

There is abundant evidence that economic incentives can influence the productivity and interests of talented individuals who potentially can make a huge contribution to the accumulation of wealth. For individuals or groups of individuals to have an incentive to adopt more advanced technology or engage in the creation of new ideas requires an institutional framework which allows for an adequate rate of return. In an interesting development of Schumpeter's theory of entrepreneurship Baumol (1990) has shown that by extending the model, so that it encompasses the 'allocation' of entrepreneurial skills, the power of the model to yield policy insights is greatly enhanced. In Schumpeter's (1934) analysis he identifies five forms of entrepreneurial activity in addition to those related to fostering improvements in technology, namely:

1. the introduction of new goods and/or new quality of an existing good;
2. the introduction of a new production method;
3. the opening of a new market;
4. the 'conquest' of a new source of supply of raw materials;
5. the new organization of any industry.

Baumol (1990) argues that Schumpeter's list is deficient and in need of extension to include:

6. 'innovative acts of technology transfer that take advantage of opportunities to introduce already available technology to geographical locales whose suitability for the purpose had previously gone unrecognised or at least unused';
7. 'innovations in rent seeking procedures'.

This last category is of crucial importance because it includes what Baumol calls 'acts of unproductive entrepreneurship'. Given that Baumol defines entrepreneurs as 'persons who are ingenious and creative in finding ways that add to their own wealth, power, and prestige', it follows that one of the main determinants of the allocation of entrepreneurial talent at a particular time and place will be the 'prevailing rules of the game that govern the payoff of one entrepreneurial activity relative to another'. Talented individuals are naturally attracted to activities with the highest private returns. There is no guarantee that such activities will always have the highest social rate of return. There-

fore entrepreneurs can be 'unproductive', even 'destructive', as well as productive from society's point of view. While in some economies talented people become conventional business entrepreneurs, in others talent is attracted to the government bureaucracy, the armed forces, religion, crime and other rent-seeking activities. Since the 'rules of the game' can and do change, we can expect to see a reallocation of entrepreneurial talent appropriate to any new environment. The forms of entrepreneurial behaviour that we observe will obviously change over historical time and across geographical space. Therefore the allocation of entrepreneurship between productive, unproductive and destructive activities cannot but have a 'profound effect' on the innovativeness, and hence growth, of any economy. For example, Landes (1969) suggests that the reason why the Industrial Revolution began in England rather than in France is linked to the allocation of talent. Any country interested in growth must ensure that its most able people are allocated to the productive sectors of the economy. Murphy et al. (1991) argue that a possible reason for the productivity growth slowdown in the US economy during the 1970s could be the misallocation of human capital because talented individuals have increasingly become rent seekers (for example lawyers) rather than producers (for example engineers). They conclude that the 'allocation of talented people to entrepreneurship is good for growth and their allocation to rent seeking is bad for growth'. In Baumol's view the predominant form of unproductive entrepreneurship in economies today is rent seeking, and the prevailing laws, regulations and structure of financial incentives will inevitably have a major effect on the 'allocation of talent'.

These insights suggest that a fruitful line of research is to focus on cross-country differences in incentive structures facing entrepreneurs with respect to encouragement to create new enterprises, adopt new technologies and thereby increase growth. If barriers to productive entrepreneurship are deliberately created by specific groups who have a clear vested interest in the *status quo*, then the task of economists is to offer policy advice about how to design and establish institutions that minimize this 'unproductive' behaviour. When an environment is created that is conducive to the adoption of new ideas by entrepreneurs, a type of capital is generated which economists have referred to as 'organizational' or 'business' capital, which exists independently of the entrepreneur. Countries that lack organizational capital will remain unattractive to foreign direct investment.

Trust between economic agents is a crucial determinant of the cost of transactions. This idea has a long pedigree (Fukuyama, 1995). For example, John Stuart Mill (1848) noted that there are counties in Europe

> where the most serious impediment to conducting business concerns on a large scale, is the rarity of persons who are supposed fit to be trusted with the receipt

> and expenditure of large sums of money … The advantage to mankind of being able to trust one another, penetrates into every crevice and cranny of human life: the economical is perhaps the smallest part of it, yet even this is incalculable.

In a recent paper, Zak and Knack (2001) have taken up this insight and show that the extent of trust in an economy 'significantly' influences growth rates, and that 'high trust societies produce more output than low trust societies'. In economies where there is a high level of trust between transactors, the rate of investment and economic growth is likely to be higher than in low-trust environments. This finding supports the earlier empirical research of Knack and Keefer (1995, 1997a, 1997b), who find a positive relationship between trust and growth for a sample of 29 market economies. Zak and Knack argue that trust is lower in countries where: (i) there is an absence of formal (laws, contract enforcement) and informal (ostracism, guilt, loss of reputation) mechanisms and institutions which deter and punish cheaters and constrain opportunistic behaviour; (ii) population heterogeneity (ethnic diversity) is greater; and (iii) inequalities are more pervasive. Easterly and Levine (1997) find that ethnic diversity in Africa reduces the rate of economic growth since diverse groups find it more difficult to reach cooperative solutions and scarce resources are wasted because of continuous distributional struggles, of which civil war, 'ethnic cleansing' and genocide are the most extreme manifestations (Bosnia, Rwanda, Kosovo, Afghanistan). Collier's (2001) research suggests that ethnically diverse societies are 'peculiarly ill suited to dictatorship' and that providing there is not 'ethnic dominance' in the political system, then democratic institutions can greatly reduce the potential adverse economic impact of ethnic diversity and the wars of attrition that can take place between competing groups. Easterly (2001b) argues that formal institutions that protect minorities and guarantee freedom from expropriation and contract repudiation can 'constrain the amount of damage that one ethnic group could do to another'. In Easterly's framework of analysis the following relationship holds:

$$\text{Ethnic conflict} = f(\text{ethnic diversity, institutional quality})$$

Easterly's research findings show that ethnic diversity does not lower growth or result in worse economic policies providing that good institutions are in place (Snowdon, 2003a). Good institutions also 'lower the risk of wars and genocides that might otherwise result from ethnic fractionalisation'. Rodrik (1999a, 1999b, 2000) has also shown how societies with deep social divisions and a lack of democratic institutions of conflict management are highly vulnerable to exogenous economic shocks (see also Alesina and Rodrik, 1994). The adverse effect of an external shock (S) on economic growth is the

bigger the greater the latent social conflict (*LSC*) and the weaker are a society's institutions of conflict management (*ICM*). Rodrik's hypothesis can be captured by the following relationship:

$$\Delta \text{ growth} = -S(LSC/ICM)$$

Rodrik (1999a) uses this framework to explain the numerous 'growth collapses' that occurred across the world economy after the economic shocks of the 1970s. In his empirical analysis countries with democratic and high-quality government institutions demonstrated better macroeconomic management and as a result experienced less volatility in their growth rates than countries with weak institutions of conflict management.

11.19 Trade and Economic Growth

Another important strand of the growth literature points to the importance of increasing international economic integration ('globalization') as a major fundamental determinant of economic growth. It is argued that there are sound theoretical reasons for believing that more open economies grow faster than more closed economies. In a recent survey of the literature, Lewer and Van den Berg (2003) conclude that the 'impact of trade on economic growth appears to be very important for human welfare'. Moreover, the proponents of this view support their case with numerous empirical studies and conclude that, overall, 'globalization' has had a positive effect on economic growth (Sachs and Warner, 1995; Krueger, 1997; Edwards, 1998; Frankel and Romer, 1999; O'Rourke and Williamsom, 1999; Baldwin, 2000; Bhagwati and Srinivasan, 2002; Dollar and Kraay, 2003, 2004; Bordo et al., 2002b; Bhagwati, 2004; Winters, 2004; see also Snowdon, 2001c, 2003c).

While economists recognize that freer trade will have a 'level effect' in raising a country's output once and for all, the likely impact on the rate of growth of output is much more controversial. For example, Lucas (1988) suggests that the removal of barriers to trade may induce a series of boosts to output which are level effects disguised as growth effects. In the standard open economy version of the Solow growth model, trade liberalization can have a temporary, but no permanent, effect on the long-run rate of growth. As Rivera-Batiz and Romer (1991) note, the main problem is that economists do not have a 'rigorous model' to justify their belief that increased economic integration will tend to increase the long-run rate of growth. Rodriguez and Rodrik (2000) have claimed that little evidence can be found 'that open trade policies – in the sense of lower tariff and non-tariff barriers to trade – are significantly associated with economic growth'. In response, Bhagwati and Srinivasan (2001) argue that an open trading regime not only allows the

beneficial level effects to be gained from comparative advantage but also has a positive influence on growth. But rather than relying on 'endless' (and 'mindless') cross-county regression evidence, Bhagwati and Srinivasan appeal to the numerous in-depth country case studies which support the growth-inducing effects of greater openness (Bhagwati, 1978; Krueger, 1978; Balassa, 1989; Edwards, 1998). Krueger (1997, 1998) is in no doubt that the developing countries that followed more outward-oriented strategies have grown faster on a sustained basis than those who 'blithely abandoned' the principle of comparative advantage and adopted, and maintained for long periods, import substitution (ISI) policies. Krueger links the initial hostility to outward-oriented strategies after 1945 to the powerful influence of *dirigiste* ideas which in turn were a legacy of the Great Depression and the apparent success of state-led development in the Soviet Union. Bhagwati (1993) recalls how India was propelled towards a 'harmful' ISI strategy because many of India's influential economists during the 1950s, including P.C. Mahalanobis, took the idea of 'elasticity pessimism' with respect to exports too seriously. In contrast to India, the experience of over four decades of outstanding economic growth performance achieved by the East Asian 'miracle' economies is also positively linked to their choice of open trade regime. The empirical work of Sachs and Warner (1995), Ben-David (1996), Edwards (1993, 1998), and Ben-David and Loewy (1998) adds support to the mainstream view that there is a strong link between free trade and income convergence among nations. Lawrence and Weinstein (2001), while providing support to those who advocate more liberal trade policies, reject the 'export fetishism' of some earlier studies and emphasize instead the growth-enhancing effects of an increasing share of imports in GDP via its effect on innovation and learning (see Rodrik, 1995). What is certain is that a more open economy will have access to cheaper imported capital goods from the world market (see DeLong and Summers, 1993; Jones, 1994).

An important factor influencing convergence identified by Sachs and Warner (1995) is the degree of openness of an economy:

> We suggest that the most parsimonious reading of the evidence is that convergence can be achieved by all countries, even those with low initial level of skills, as long as they are open and integrated in the world economy … the convergence club is the club of economies linked together by international trade … In terms of the conditional convergence hypothesis, we argue that the apparent differences in long-term income levels are not differences due to fundamental tastes and technologies, but rather to policies regarding economic integration.

Sachs and Warner identify a 'closed' trading regime as one that has at least one of the following characteristics: (i) non tariff barriers covering 40 per cent or more of trade; (ii) average tariff rates of 40 per cent or more; (iii) a

black market exchange rate that is depreciated by 20 per cent or more relative to the official exchange rate, on average, during the 1970s and 1980s; (iv) a socialist economic system (as defined by Kornai, 1992); and (v) a state monopoly on major exports. An open economy is one where none of the above conditions hold. The case of China is the only real puzzle. However, as Sachs and Warner explain, with respect to international trade, the Chinese economy was essentially liberalized for non-state firms, especially those operating in the Special Economic Zones in the coastal areas. And the fastest-growing areas in China in the period 1978–94 were all coastal provinces with the exception of Xinjiang (Ying, 1999).

The research findings of Sachs and Warner lead them to four important conclusions: first, 'there is strong evidence of unconditional convergence for open countries, and no evidence of unconditional convergence for closed countries'; second, 'closed countries systematically grow more slowly than do open countries, showing that "good" policies matter'; third, 'the role of trade policy continues after controlling for other growth factors'; and fourth, 'poor trade policies seem to affect growth directly, controlling for other factors, and to affect the rate of accumulation of physical capital'. Therefore, Sachs and Warner's reading of the evidence suggests that trade policy should be viewed as 'the primary instrument of reform' where trade policy serves as a proxy for an entire array of market reforms.

At the theoretical level, work in the endogenous growth literature has emphasized the importance of the flow of ideas in stimulating technological innovation (P. Romer, 1990, 1993, 1994b). In this context the greater a country's exposure to the world outside the more it is likely to gain from the research and development activities of other countries, including new ideas relating to organizational methods. As Paul Romer argues, 'The key role for trade is that it lets developing countries get access to ideas that exist in the rest of the world'. This view also receives support from Robert Solow, who suggests that 'The only way you can make sense of trade having an effect on the long-run growth rate is not so much whether the country is export led, but whether the country is in contact with the rest of the world' (see Solow interview at the end of this chapter). Edwards (1998) provides a simple framework for considering the relationship between TFP growth and openness. This framework is summarized in equations (11.44) and (11.45).

$$Y_t = A_t f(K_t, L_t) \tag{11.44}$$

$$\dot{A}/A = \theta + \omega(W - A)/A \tag{11.45}$$

Here Y_t is GDP, A_t is the stock of knowledge, or TFP, K_t is physical capital and L_t is labour measured in efficiency units. Growth will depend on the rate

of change of A_t, K_t and L_t. Edwards assumes that there are two sources of TFP growth: first, a domestic source fuelled by innovation and dependent on domestic human capital (education); and second, an international source 'related to the rate at which the country is able to absorb (or imitate) technological progress originating in the leading nations'. Imitation depends on a 'catch-up' term. Those countries furthest from the technological frontier have the greatest potential for imitation. In equation (11.45) $\dot{A}/A$ is the rate of growth of TFP, θ is the domestic rate of innovation, ω is the speed at which a country is able to close its knowledge gap and is influenced by trade policies, W is the world stock of knowledge assumed to grow at a rate g (where $g \geq \theta$). For the world's technological leader (the USA since the 1890s) $g = \theta$ and $W = A$. Edwards argues that, in keeping with many ideas-based models of growth, more open economies 'have a greater ability to absorb ideas from the rest of the world and, thus, have a higher ω' and 'an important property of this simple model is that countries that liberalise trade will experience transitional productivity growth that exceeds that of countries that maintain their trade distortions'. Because trade between nations is likely to act as a conduit for the dissemination of ideas, inward-looking trade and development strategies which erect barriers to trade will therefore inhibit the transmission of knowledge.

The extent of North–South R&D spillovers has been investigated by Coe et al. (1997). Given that almost all R&D activity is carried out in the developed countries, there are clearly opportunities for developing countries to benefit from knowledge spillovers, especially from the USA. The empirical evidence presented by Coe et al. indicates that the total factor productivity of developing countries is 'positively and significantly related to R & D in their industrial country trade partners and to their imports of machinery and equipment from the industrial countries'.

In the model developed by Ben-David and Loewy (1998) openness creates greater competitive pressure on domestic firms which, in response, seek to acquire foreign knowledge relating to production processes and techniques (see also Parente and Prescott, 2000; Baumol, 2002). Hence trade flows facilitate the transfer of ideas and stimulate the growth of the economy. Because many poor countries have adopted protectionist strategies, their trade barriers act as a 'buffer that limits knowledge spillovers'. Ben-David and Loewy argue that so long as such barriers persist, the income gap between countries will continue to exist.

Remarkably, Adam Smith (1776) anticipated the argument that free trade facilitated the flow of ideas and knowledge:

> Nothing seems more likely to establish this equality of force than the mutual communication of knowledge and of all sorts of improvements which an extensive

> commerce from all countries to all countries naturally, or rather necessarily, carries with it.

In Irwin's (1996) view, these insights of Smith, relating to technology transfer and the dynamic benefits of openness, 'were outstanding for the period in which he was writing'.

While the majority view of economists leans towards the Bhagwati position on trade and growth, Rodrik (1995, 1999a) remains critical of those who place too much emphasis on 'globalization' and exports as the easy road to economic development. He continues to emphasize that policy makers in developing countries need to formulate a growth strategy that recognizes the importance of domestic institutions (including democracy) and domestic investors rather than one that is built solely around a 'hazardous obsession with global integration'. Unlike Bhagwati and Krueger, who have provided a sustained critique of ISI policies, Rodrik (1999a, 1999b) argues that the wrong lessons have been learned about the growth experiences of countries that adopted ISI policies. According to Rodrik, ISI policies worked very well for about twenty years before the late 1970s, and the subsequent growth collapses and disappointing economic performance of many ISI-adopting countries 'had little to do with ISI policies'. The developing countries that successfully weathered the economic shocks of the 1970s and 1980s were those that quickly and decisively made the appropriate macroeconomic adjustments and also had effective domestic institutions of conflict management. So the reason that the East Asian economies have been better equipped to cope with economic turbulence compared to Latin America and sub-Saharan Africa is not that they were outward-oriented and the others remained closed; rather it was because Latin America and sub-Saharan Africa 'did a much worse job' in managing and absorbing the shocks. Rodrik does not deny the potential benefits of openness, but he warns policy makers and fellow economists that for openness to be a successful component of any development strategy, it must not be regarded as a substitute for a domestic strategy. The 'knee-jerk' globalizers among policy makers mistakenly seem to imply that globalization, by itself, can work miracles for a developing country's economy. Therefore Rodrik remains a strong advocate of the importance of domestic investment, human and physical capital formation, as the fundamental determinant of economic growth. While empirically the relationship between investment and growth 'tends to be erratic in the short run … cross-national studies have shown that investment is one of the few robust correlates of economic growth over horizons spanning decades' (see Rodrik, 1995, 1999a; see also Vamvakidis, 2002).

A reasonable conclusion from the above discussion is that it appears to be openness rather than exports *per se* which seems to matter for enhanced

economic performance, but also that to fully benefit from the potential benefits of openness developing counties need a complementary domestic strategy that includes building 'institutions for high-quality growth' (Rodrik, 2000, 2003, 2005) . Openness is a means to an end, not an end in itself. Interestingly, as the research of Alesina and Spolare (2003) shows, as the world economy becomes more open, the size of a country's domestic market becomes less important as a positive influence on the level and growth of productivity. As globalization spreads, the benefits of size diminish. Predictably the number of countries in the world has increased from 74 in 1945 to 192 in 2004.

11.20 Geography and Growth

In recent years several scholars have revived the idea that geography has an important influence on economic performance. There have been two strands in this literature. The first, represented by the work of economists such as Paul Krugman, Anthony Venables and Michael Porter, highlights the role of increasing returns, agglomeration, size, clusters and location in the productivity performance of nations and regions (see Krugman, 1991a, 1991b, 1997; Krugman and Venables, 1995; Martin, 1999; Crafts and Venables, 2002; Porter, 2003; Yang, 2003). With intellectual roots in the work of Alwyn Young, Gunnar Myrdal and Nicholas Kaldor, the 'new economic geography' models highlight the impact of cumulative causation effects whereby success breeds success. In these models globalization can initiate cumulative processes that lead to the persistence of uneven spatial (urban, regional, and international) development. As Crafts and Venables (2002) point out, in a world dominated by increasing returns, cumulative causation, agglomerations effects and path dependency, the prospects that increasing international integration will lead to convergence are much less certain.

A second strand in the literature emphasizes the direct impact that geography can have through climate, natural resources and topography. Such factors obviously influence the health of a population, agricultural productivity, the economic structure of an economy, transport costs and the diffusion of information and knowledge. Geography, it is argued, plays an important role in determining the level and growth of income per capita (see Diamond, 1997; Bloom and Sachs, 1998; Gallup et al., 1998; Bloom et al., 2003). For a critique of this literature see Acemoglu et al., 2001, 2002a).

An important stimulus to the revival of interest in the impact of geography on economic performance comes from the increasing recognition that income per capita and latitude are closely related. Countries nearer to the equator, with a few exceptions (such as Singapore), have lower income per capita and HDI scores than countries located in more temperate zones. The strong

negative empirical association between living standards and proximity to the tropical latitudes is strongly influenced by the 'dismal growth performance of the African continent' which has produced the 'worst economic disaster of the twentieth century' (Artadi and Sala-i-Martin, 2003; see also Easterly and Levine, 1997; Collier and Gunning, 1999a, 1999b; Herbst, 2000). What accounts for the extraordinarily poor economic performance of sub-Saharan African economies during the second half of the twentieth century, particularly since decolonization?

Bloom and Sachs (1998) have argued that six sets of factors have featured in various accounts of the poor economic performance of sub-Saharan African economies, namely explanations based on:

1. unfavourable external factors related to colonial and cold war legacies;
2. volatility in primary exports terms of trade;
3. internal politics conducive to authoritarianism and corruption;
4. *dirigiste* economic policies emphasizing import substitution and fiscal profligacy;
5. demographic trends involving rapid population growth and a 'stalled demographic transition'; and
6. ethnic diversity and low levels of social capital (trust).

However, in addition to these factors, which have all played some role, Bloom and Sachs believe that economists ought to 'lift their gaze above macroeconomic policies and market liberalisation' and recognize the constraining influence on sub-Saharan Africa development of its 'extraordinarily disadvantageous geography'. By having the highest proportion of land area (93 per cent) and population (659 million, in 2000) of all the world's tropical regions, sub-Saharan Africa, by virtue of its climate, soils, topography and disease ecology, suffers from low agricultural productivity, poor integration with the international economy and poor health and high disease burdens now boosted with the onset of an AIDS epidemic. According to the World Bank Development Report (2002), sub-Saharan Africa has the lowest per capita income of all the major regions of the world ($480, and PPP$1560). In 1999 sub-Saharan Africa's life expectancy of 47 years was also the lowest, and sub-Saharan Africa's under-5 mortality rate of 159 per 1000 births the highest, in the world. Certainly, the evidence supports a positive link between the health of nations and their ability to accumulate wealth (Bloom et al., 2004).

The influence of geographical factors on economic growth and development was not lost on Adam Smith (1776), who recognized that success in trade was greatly enhanced by having easy access to water transportation. Smith (1776) noted that

> it is upon the sea-coast, and along the river banks of navigable rivers, that industry of every kind naturally begins to subdivide and improve itself … All the inland parts of Africa, and all that part of Asia which lies any considerable way north of the Black and Caspian seas … seem in all ages of the world to have been in the same barbarous and uncivilised state in which we find them at present.

Recently, Rappaport and Sachs (2003) have shown that economic activity in the USA is overwhelmingly concentrated along or near its ocean and Great Lakes coastal regions. As Adam Smith recognized, proximity to coastal regions greatly enhances productivity performance and the quality of life.

Whilst not arguing a new case of geographical determinism, and also recognizing the crucial role played by economic policies, Bloom and Sachs believe that 'good policies must be tailored to geographical realities'. They conclude that Africa will be well served if economists take advantage in their research of 'much greater cross-fertilisation' from the accumulated knowledge in other fields such as demography, epidemiology, agronomy, ecology and geography. Thus an important divide in the world does exist, but it is not between North and South; rather it is between countries located in temperate latitudes compared to those in the tropics (see also Sachs and Warner, 1997; Diamond, 1997; Hall and Jones, 1999; Landes, 1990, 1998; Sachs, 2003).

Finally in this section we draw attention to the recent research on the 'natural resource curse', that is, the tendency of some countries that possess abundant natural resources to grow more slowly than natural resource-poor countries. While democracies such as the USA, Canada and Norway tend to manage oil and other natural resources well, this is far from the case in countries governed by predatory kleptocratic autocrats where the presence of 'black gold' stimulates rent-seeking behaviour, political instability and, in the extreme, civil war (see Sachs and Warner, 2001; Eifert et al., 2003). As Sala-i-Martin and Subramanian (2003) document, 'Nigeria has been a disastrous development experience' despite having large oil resources. Successive corrupt military dictatorships have simply plundered the oil revenues. In contrast to Nigeria, the experience of Botswana, with its lucrative diamond resources, has been completely different. The economic success of Botswana is mainly due to the quality of its governance and institutions (see Acemoglu et al., 2003).

11.21 Growth in History: In Search of a Unified Theory

Before considering the 'ideal conditions' for economic growth we will briefly survey the new and exciting literature that attempts to provide a unified account of the 'Great Escape' of leading world economies from 'Malthusian stagnation' to a regime of 'modern economic growth'. Prominent here has been the research of scholars such as Daron Acemoglu, Gregory Clark, Richard Easterlin, Oded Galor, Martin Goodfriend, Gary Hansen, Simon Johnson,

Charles Jones, Eric Jones, Michael Kremer, David Landes, Robert Lucas, Angus Maddison, John McDermott, Omer Moav, Joel Mokyr, Douglass North, Stephen Parente, Kenneth Pomeranz, Edward Prescott, James Robinson and David Weil. The main puzzles to explain are these:

1. Why did no country or region of the world, before the eighteenth century (the Malthusian era), experience lasting intensive growth, that is, sustained increases in per capita GDP?
2. What led to the 'Industrial Revolution', and was this 'Revolution' inevitable?
3. What caused the 'Great Divergence' in living standards across the world that has taken place during the last 250 years?

It is difficult to imagine that there are any bigger or more difficult questions for economists to answer! At the moment, scholars working on providing answers to these questions are at the frontiers of growth theory and empirics. Increasingly, in order to provide plausible answers to such questions, growth theorists have begun to appreciate the value of economic history and the previous research carried out by economic historians. Indeed, there has recently emerged a growing synergy between growth theory, economic history and development economics that has been long overdue. While the Solow and Romer models may provide convincing explanations of the modern experience of economic growth in developed economies, they do not account for the major growth transition that occurred with the onset of the Industrial Revolution. What is needed is a unified theory that can account for the major features of the Malthusian era as well as the modern growth regime as documented in Table 11.1. Currently there are several plausible accounts or 'stories' of the evolution of world living standards.

The natural selection, evolution story

Galor and Moav (2001, 2002) highlight the interplay between Darwinian and Malthusian forces. The struggle for survival during the epoch of Malthusian stagnation gradually leads to an evolutionary improvement in the quality of human capital, which in turn stimulates an increase in the rate of technological progress. This eventually creates the conditions conducive to a 'take-off' of sustained economic growth. Galor and Weil (1999, 2000) carry this story forward and explain how an endogenous transition takes place with an economy moving from a 'Malthusian Regime' to a 'Post-Mathusian Regime' before finally entering the 'Modern Growth Regime'. The first two regimes are separated by an acceleration of technological progress while the latter two are separated by the demographic transition which is driven by utility-maximizing fertility behaviour (see Lee, 2003).

The population, ideas and property rights story
Michael Kremer (1993), building on the work of Paul Romer, constructs a model where technological progress is driven by ideas. In turn the number of ideas depends on the size of the population. Therefore, during the Malthusian era, while improvements in living standards are negligible, technological progress causes the size of the population to increase, which further stimulates technological progress through the creation of more ideas. Kremer's model predicts that, historically, the growth rate of population will be proportional to its level, at least before the worldwide spread of the demographic transition in the latter part of the twentieth century (see Lee, 2003). Charles Jones (2001b) adds to this story a key requirement that in order for technological progress to win the race against Malthusian diminishing returns, not only do there need to be increasing returns to accumulable factors; there also need to develop 'innovation-promoting institutions' as emphasized by North (1990).

The Mathus to Solow story
Hansen and Prescott (2002) and Parente and Prescott (2005) build models where an economy is initially dominated by a land-intensive, low-productivity 'Malthus technology' with low knowledge input. Eventually, as knowledge grows, driven by the profit motive, the economy gradually switches to one that is dominated by a much more productive 'Solow technology'. An earlier model of Goodfriend and McDermott (1995) emphasizes the transition from household to market production driven by the increasing returns to specialization made possible by a growing population.

The institutions and property rights story
North and Weingast (1989) highlight the positive impact that the establishment of more secure property rights had on innovation and entrepreneurship in Britain before the Industrial Revolution. Acemoglu et al. (2002b), building on this idea, trace the rise of the colonial 'Atlantic trader' economies after 1500 and link their subsequent growth success to the influence of the commercial bourgeoisie who demanded and obtained changes in institutions that led to greater protection of property rights. This is turn provided a foundation for the Industrial Revolution to take place in Britain.

The 'Gifts of Athena' story
Mokyr (1990, 2002, 2005) surveys the history of technological change and traces the intellectual roots of the Industrial Revolution to important changes in the method and culture involved with the creation and dissemination of new knowledge. 'The Scientific method that evolved in seventeenth century Western Europe meant that observation and experience were placed in the

public domain' (Mokyr, 2005) and scientific knowledge became a public good. 'Open science' and verification, rewarded by fame and recognition, became part of what Mokyr calls the 'Industrial Enlightenment'. The notion that economic progress was possible dominated this new enlightenment. The wave of 'macroinventions' and 'microinventions' (the 'wave of gadgets') that characterized the Industrial Revolution would not have been possible without these intellectual roots. As Mokyr notes, 'knowledge creates opportunities, but it does not guarantee action' (2005). Moreover, the 'emphasis on the Enlightenment illustrates how economists should think about culture and cultural beliefs'.

These are just some of the recent stories that attempt to provide a unified account of the evolution of world income over the very long run. The interested reader should also consult Jones (1988), Easterlin (1996), Landes (1998), Lucas (2002), Maddison (2001) and Clark (2003).

11.22 The Ideal Conditions for Growth and Development: Rediscovering Old Truths

Given our extended discussion of economic growth, it seems appropriate to conclude by asking the question, 'what are the ideal conditions likely to foster significant improvements in living standards?' In order to generate the 'ideal conditions' for material progress, Landes (1998) has argued that the lessons of history imply the following prerequisites:

1. an environment which fosters initiative, competition and emulation;
2. that job selection be based on merit, competence and performance;
3. financial rewards should relate to effort/enterprise;
4. the economy must be fully exposed to existing technological knowledge;
5. that education be widespread amongst the population.

These prerequisites in turn imply:

6. an absence of discrimination based on irrelevant criteria (race, religion, gender);
7. an expressed preference for scientific rationality over 'magic and superstition'.

The political and social institutions conducive to achieving these goals will involve:

8. providing laws protecting the security of property, personal liberty and secure rights of contract;
9. reducing rent-seeking behaviour;
10. providing for a more stable, moderate, honest, efficient and responsive (democratic) government operating within a framework of publicly known rules (Tanzi, 1999).

The importance of these preconditions for sustainable growth has been increasingly recognized, although frequently deliberately ignored, since the publication of Adam Smith's *Wealth of Nations* in 1776. While no economy in the world meets Landes's 'ideal' conditions for material progress, it is obvious that some come much closer to meeting the above criteria than others.

Along similar lines, Rodrik (2005) identifies certain desirable 'meta principles' that seem to apply across the globe regardless of history, geography and stage of development. They include the importance of: incentives; security of property rights; contract enforcement and the rule of law; the power of competition; hard budget constraints; macroeconomic and financial stability (low inflation, prudent regulation and fiscal sustainability); and targeted redistribution that minimizes distortions to incentives.

To a large extent these 'meta principles' are 'institution free' in the sense that they do not imply any fixed set of ideas about appropriate institutional arrangements. In other words, there are many possible models of a mixed economy and in practice the capitalist market-oriented economies that we observe across the world exhibit a 'diverse range of institutional arrangements'. Every capitalist system consists of an amalgam of public and private institutions and, as the 'new comparative economics' recognizes, European, Japanese, East Asian and American forms of capitalism differ. However, as Rodrik argues, in each case there are firmly in place 'market-sustaining institutions' which he divides into four categories:

1. market-creating institutions (property rights and contract enforcement);
2. market-regulating institutions (regulatory authorities);
3. market-stabilizing institutions (monetary, fiscal and financial authorities);
4. market-legitimizing institutions (democracy and social protection).

Given this framework, Rodrik concludes that there is 'no unique correspondence between the *functions* that good institutions perform and the *form* that such institutions take'. There is plenty of room for 'institutional diversity' consistent with meeting the broad 'meta principles' (see Snowdon, 2002d).

On the basis of the above discussion many developing and transition economies have a long way to go before they can hope to transform their economies

into systems that are capable of generating sustained improvements in living standards. Without a significant change of direction what hope is there for the citizens of countries such as North Korea? Thus while growth theory and empirical research show that poor countries have enormous potential for catch-up and convergence, these advantages will fail to generate positive results on growth in countries with an inadequate political, legal and regulatory framework. Research also indicates that economic growth does in general benefit the poorest groups in society, so anyone who claims to care about the poor should also favour 'the growth-enhancing policies of good rule of law, fiscal discipline, and openness to international trade' (Dollar and Kraay, 2002b).

As we enter the new millennium liberal democracy is on the increase across the nations of the world (Barro, 1997; World Bank, 1997). Whether this 'world-wide liberal revolution' constitutes the 'end point of mankind's ideological evolution', as argued by Fukuyama (1989, 1992), is debatable. Huntington (1996) offers a more pessimistic scenario with his controversial argument that the twenty-first century will be characterized by 'The Clash of Civilisations' based on cultural–religious divides! However, if the trend towards increasing democratization continues, and the pessimistic predictions of Huntington turn out to be wrong, the prospects for peace and 'spreading the wealth' in the twenty-first century are greatly increased. Research indicates that while democracies are just as likely to fight wars in general, and are also more likely to prevail in conflicts with autocratic regimes, they are also much less likely to wage war with one another. Compared to dictatorships, democratic governments are more accountable to their electorates and have better institutional means of conflict management and resolution (Lake, 1992; Dixon, 1994). If the twenty-first century turns out to be more democratic and peaceful than the twentieth century, then we can be much more optimistic about the prospects for growth and poverty reduction across all the world's economies. As suggested by Minier (1998), a useful motto for the twenty-first century is '*liberté, egalité, fraternité and prosperité*'. In such a world 'Convergence, Big Time' is not impossible.

ROBERT M. SOLOW

Robert Solow was born in 1924 in Brooklyn, New York and obtained his BA, MA and PhD from Harvard University in 1947, 1949 and 1951 respectively. He began his academic career as an Assistant Professor of Statistics (1950–54) at the Massachusetts Institute of Technology (MIT), before becoming Assistant Professor of Economics (1954–8), Professor of Economics (1958–73) and Institute Professor of Economics (1973–95) at MIT. Since 1995 he has been Institute Professor Emeritus of Economics at MIT.

Professor Solow is best known for his seminal work on growth theory and capital theory, and for his development and championing of neo-Keynesian economics. In 1987 he was awarded the Nobel Memorial Prize in Economics 'for his contributions to the theory of economic growth'. Among his best-known books are: *Linear Programming and Economic Analysis* (McGraw-Hill, 1958), co-authored with Robert Dorfman and Paul Samuelson; *Capital Theory and the Rate of Return* (North-Holland, 1963); *Inflation, Unemployment and Monetary Policy* (MIT Press, 1998), co-authored with John Taylor; and *Growth Theory: An Exposition* (2nd edn, Oxford University Press, 2000). His most widely read articles include: 'A Contribution to the Theory of Economic Growth', *Quarterly Journal of Economics* (1956); 'Technical Change and the Aggregate Production Function', *Review of Economics and Statistics* (1957); 'Analytical Aspects of Anti-Inflation Policy', *American Economic Review* (1960), co-authored with Paul Samuelson; 'Does Fiscal Policy Matter?' *Jour-*

nal of Public Economics (1973), co-authored with Alan Blinder; and 'On Theories of Unemployment', *American Economic Review* (1980).

We interviewed Professor Solow in Chicago, in his hotel room, on 4 January 1998, while attending the annual conference of the American Economic Association. In what follows we present an abridged version of the interview, which focuses on Professor Solow's views on economic growth. The full text of the interview, including Professor Solow's views on the development of modern macroeconomics, can be found in Snowdon and Vane (1999b).

Background information

When did you first decide to study economics?
Well, there is a story to that. I came to Harvard College in 1940 as a 16-year-old freshman with no intention of studying economics; I did not even know what economics was. At that point I thought I might be a biologist but I proved to be no good at that so I started off as a major in general social science. I studied subjects like elementary economics, psychology, sociology and anthropology. The reason I was interested in social science was just the circumstances of the time. Remember it was 1940, the Depression was just over, and the war had just begun. In 1942, after two years, I quit Harvard College and joined the army, which seemed more important to me then. In 1945 I returned to education and I said to the girl I left behind and who has been my wife ever since, 'You majored in economics; was it interesting?' When she said yes I decided to give it a try. At the time I was under pressure to choose something to study because I was discharged in August and the school term was due to start in September. I was still an undergraduate. Anyway, it turned out all right. So the reason I studied economics was related both to my general interest in what was happening – why society was not working so well in the 1930s and 1940s – and to sheer desperation because I had to do something in a hurry.

As a student, which of your teachers inspired your interest in economics?
Mainly Wassily Leontief, who taught me for one course, even before I joined the army. In those days Harvard College had a tutorial system and every student majoring in economics had a member of the faculty assigned to him as a tutor. We met once a week and it was obviously an imitation of the Oxford and Cambridge system. Wassily was my tutor and I really learnt my economics from him; he was undoubtedly the main person who inspired my interest in economics. The only other teacher in those days who really caught my imagination was Dick Goodwin, who had been my teacher in the elementary economics course that I had taken in 1940–41. I hit it off with

him very well. After the war when I came back I studied more economics with him.

Which economists have had the greatest influence on the direction that your own work has taken?
Since I completed my PhD degree, Paul Samuelson and Jim Tobin – both very good friends – are the people whose way of doing economics I admired and still admire. They were representatives of what I now (I did not see it then) think of as the new style of doing economics after the war. Economics went from being a sort of cultural subject to a model-building subject, and I liked that. Paul Samuelson and Jim Tobin were the people who for me exemplified that new approach. The other name I would mention, but not from a personal contact point of view, only from his work, was Lloyd Metzler. I read Metzler's work after I had read Samuelson's [1939] multiplier–accelerator papers. Metzler's [1941, 1947] papers on inventory cycles and 'Wealth, Saving and the Rate of Interest' [1951] were absolutely splendid. I did not know Lloyd Metzler very well because he had gone off to Chicago by then and later he suffered a terrible brain tumour. After that he was no longer the real Lloyd Metzler.

Economic growth

The whole issue of economic growth has seen a regeneration of interest in recent years and many prominent economists like Robert Barro and Xavier Sala-i-Martin [1995] regard it as the part of macroeconomics that really matters. There seems to have been a neglect of Nobel Prize awards in this area, so do you anticipate that this relative neglect will be rectified in the future given the importance of growth for human welfare?
Well, I would not describe it as a relative neglect. I think that what looks like neglect is actually something quite different. They started the Nobel Prize awards in Economics in 1969, and, unlike Physics or Chemistry, which had been going for years, there was a long backlog of people who were clearly of that calibre, if you have that kind of prize. Therefore it was only natural to start picking them up in turn. There are exceptions. Some people came out of order; for example Ken Arrow (deservedly) came early in the awards but even he was paired with John Hicks who was old by then. In 1987 when I got the award nobody born later than me had yet been given the prize, so in a way it is still just rolling up the carpet from the old end. My view is that if growth theory, the empirical analysis of growth, and ideas connected with them, continue to be popular, the subject will attract the best people in the profession. And yes, there surely will be more awards in this area. By the way, I do not know how you count Arthur Lewis and Ted Schultz, who were interested

in economic development – Ted Schultz in a very different way – but Arthur Lewis contributed that famous 1954 paper on 'Economic Development with Unlimited Supplies of Labour'. So I would not say that there has been a neglect; I would say the timing has been natural. There are likely to be more surprises, coming at a slightly greater rate than the past, because we are now getting up to contemporary people, to economists who were doing their work fairly recently. Since 1987 there has been a real outburst of work on growth so there will be more awards in that area.

Your 1956 and 1957 papers have clearly had a profound impact on the direction of research in the area of economic growth. Can you tell us what were the main influences which led you into that research and which generated those papers?

Yes, I do recall what led me to that research. I became interested in growth for three main reasons. First, in the early 1950s everybody was interested in economic development, for the obvious reason that most of the population of the world was living in poor economies. I was passively interested in economic development, but I have never been actively interested – in a research way – in what happens in underdeveloped countries. But I got to thinking about development issues and I had read Arthur Lewis. I knew I was not going to work on development issues, but it did get me interested in the general area of economic growth. Then Paul Samuelson and I had started thinking about what later became Dorfman, Samuelson and Solow [1958], the book on *Linear Programming*. That was the second factor. In the course of that research we thought about the Von Neumann and Ramsey models. So from the optimization and linear programming end and the idea of using programming theory to deal with intertemporal optimization, I also got interested in growth. The third influence was my reading of the work of Harrod and Domar; but I guess my reaction to their ideas was a little different from some other economists. I was suspicious of the Harrod–Domar model for reasons which I have occasionally explained. It occurred to me that if the world works in the way suggested by their model, then the history of capitalism would have been much more erratic than it has been. If Harrod–Domar was a good macro model for the long run, then it is impossible to explain, to my mind, how contained economic fluctuations have been, how you can draw a trend and look at fluctuations around that trend, and how those fluctuations stay 3–4 per cent either side of trend, except for a few major depressions. I thought that there must be a way of modelling growth that does not have the knife-edge property of the Harrod–Domar model. Those were the influences which led me to the 1956 paper.

You mentioned Arthur Lewis in your answer. His model was described as a classical rather than neoclassical model. Do you think that the classical

economists made any important lasting insights on the issue of economic growth?
When you say classical economists do you mean Smith, Ricardo and Mill and so on?

Yes.
If so, that is not where I got any intellectual help, for a number of reasons. First, I am not very well read in the history of economic thought. I know the potted versions of Smith, Ricardo and Mill, but I would never trust myself to have a deep thought about classical economics. I have looked back to see if there was anything that I missed, and I would say that apart from Mill on the stationary state, and Ricardo to a certain extent as Mill's predecessor, I did not find much there other than vague ideas. They were obviously interested in the long run but that does not butter any parsnips really. The relationship of diminishing returns to the stationary state, especially in Mill, obviously has some relationship to the work I was doing in the mid-1950s. That paid off a lot. On the other hand the obvious thing on the negative side is that Ricardo at the beginning, and Mill a little later on in the course of the Industrial Revolution, were thinking about the long run and yet the notion that growth can be maintained by technological improvements did not seem to occur seriously to either of them.

Was your 1956 paper accepted for publication straight away?
Yes. I can pinpoint when I was working on it; it was in 1955. I sent it to the *Quarterly Journal of Economics* and they accepted it right away. Writing papers is very hard for me; and so throughout my whole career I have only written papers when I thought that either I had something really serious to say, or I had to produce a paper for a *Festschrift* or something like that. In the latter case anything intellectually respectable would do. But the papers that I write of my own free will are usually pretty serious, otherwise it is not worth the effort, because I really do not like doing it.

Earlier you mentioned the growth of interest in development economics which took off as a research area during the 1950s. Why did development economics emerge during this period as a separate branch of economics from growth theory?
Why did it happen that way? Well, I am going to offer a suggestion but it is not original to me. I guess it comes originally from Paul Krugman of MIT. On the whole the personality types in the profession who became interested in economic development were not model builders. They were collectors of data and generalizers from rough empirical data, like Simon Kuznets; or they were like Ted Schultz, really deeply into underdeveloped agriculture, or they

were people interested in history and backwardness for its own sake. That sort of temperament is not suited to model building. Growth theory, *par excellence*, yielded to model building. So even Arthur Lewis, whom I mentioned earlier, thought of his 1954 paper as a minor sideline to his book *The Theory of Economic Growth* [1955]. The people who got interested in the theory of economic growth were interested in model building.

When we talked to James Tobin in 1993 he remarked that the really good papers in economics always contain a surprise. Were you surprised to find that the steady state rate of growth is independent of the saving rate?
Oh yes. I wrote that up right away and wanted to publish it in spite of my dislike of writing papers. I thought it was a real shocker. It is not what I expected at all, and by the way, when I did the 1957 paper on technical change I also expected a different answer from the one that I found. I expected that the main source of growth would be capital accumulation because that is what everyone talked about and I had heard that all my life as a student. Those were both real surprises.

That 1957 paper inspired a vast literature on growth accounting, with contributions from economists such as Denison, Kendrick, Jorgenson, Maddison and others. After 40 years of work, what have we learned about the sources of economic growth?
I think we have learned a great deal, not compared with what might be learned, but compared to what we have learned in other areas of macroeconomics. The notion that technical change or the residual accounts for much more of growth than you would expect, much more of productivity increase than capital accumulation, has stood up. Where it has not stood up – as in the work of Alwyn Young [1995], Jong Il Kim and Larry Lau [1994], Sue Collins and Barry Bosworth [1996] on the four Asian Tigers – it has been fascinating and you actually learn a great deal (assuming it is all true, of course); they have recorded staggeringly rapid growth but not in the same way as the historical capitalist economies. That basic distinction between capital accumulation and the residual has proved to be very informative. We have also learned a lot about the importance of human capital, as distinct from tangible capital, but the relative importance of each is still not settled. You still find what look like perfectly sound empirical papers which come up with conflicting results about the importance of human capital, depending on the time period, the model and other factors, especially the way 'human capital' is measured. I was delighted to learn after the fact that in my 1957 paper, at the very beginning, I said that what I called technical change included a lot of things such as human capital, although I did not have that language then. But the work on growth accounting, beginning

with Edward Denison and then continuing on, has taught us a great deal about the nature of growth. I would say that the fact that the growth of the current advanced industrial economies only owes a little to the exploitation of natural resources is very interesting and this too has come out of growth accounting methods.

In 1970 the first edition of your book, Growth Theory: An Exposition, *was published. Following that, for the next 16 years, the interest of macroeconomists in the issue of economic growth, or more accurately growth theory, went into relative decline. Why do you think that happened?*
I think it happened because the profession ran out of ideas and you cannot maintain interest in any subject simply on the basis of looking more and more closely into the existing ideas. Edward Denison was still writing his books during this period, all of which I read and admired. But there were no new ideas. The merit of the contributions from Paul Romer [1986] and Bob Lucas [1988] – I do not know how to divide it up between them – is that they renewed interest in the subject by bringing in new ideas. That always attracts people to any branch of economics, and I presume the same thing is true of chemistry. So it was just a case of intellectual diminishing returns. Around 1970 we simply ran out of new ideas.

The first paper on endogenous growth in the new phase of interest in economic growth was Paul Romer's [1986] 'Increasing Returns and Long-Run Growth'. What do you think inspired the new research? Was it the convergence controversy issue which also emerged about the same time with the contributions of Abramovitz [1986] and Baumol [1986]?
Well, you are going to have to ask Lucas and Romer that question.

OK, we will ask Paul Romer that question when we interview him tomorrow.
I would have said, just from the second-hand evidence of reading their papers, that the convergence issue was more of a stimulus to Bob Lucas than it was to Paul Romer. It may have influenced Paul Romer as well, but I do not remember anything in that 1986 paper which suggests that, though I could easily have forgotten. I am inclined to think that Paul Romer had an idea, found it exciting and followed it. But Lucas gave more signs of having been fascinated by the international comparisons.

What are your views on the convergence issue? Your 1956 model predicts conditional convergence and this prediction seems to fit reasonably well to a group of countries, a 'convergence club'. Yet there are other poor countries which are showing little sign of catching up with the rich industrialized countries.

I have no independent thoughts on this at all. I just read the literature, not all of it because there is so much. But I read enough of it to develop opinions and these go roughly like this. First of all I am at heart very suspicious of all this international cross-section research. I read it, sometimes it is interesting and sometimes it is not, but in the back of my mind there is always a question as to whether I should believe it. The fundamental reason why I am dubious about it is that there is no solution to the inverse causation issue. The more right-hand-side variables that go into those regressions, the more they seem to me to be just as likely the consequences of success or failure of long-term economic growth as the cause. The second reason I am suspicious is that I learned from Ross Levine at the World Bank a long time ago that most of those results are not robust. They do not stand up if you make minor variations. The third reason I am suspicious is that I keep asking myself, do I really believe that there is a surface out there in space whose axes are labelled with all the things Robert Barro and company put on them? Do I believe that there is such a surface, and countries or points on that surface could in principle move from one place to another on it and then move back to where they began by changing their form of government or by having more or fewer assassinations? A small voice says maybe, but I would not bet anything on the existence of that surface. So I am dubious about that whole line of research. If you look at it as a pure time series problem, the way Danny Quah (1993) does, if you look at conditional convergence – and conditional convergence is the only version of this that makes sense – then the evidence does look more or less as if there really is something to the distinction between growth and development. There is a group of countries, which for one reason or another do not catch on to the railroad train as it goes by and I am inclined to attribute that to their lacking some institutional infrastructure, some sociological infrastructure, whatever. If I had to throw in my lot with one camp or the other I would support the convergence club.

Another factor which has contributed to the reawakening of interest in the growth issue has been the so-called productivity slowdown which began in the late 1960s/early 1970s. Do you believe there was a productivity slowdown and, if so, what were the possible causes?

Yes, I do believe that there was a productivity slowdown. All the debate about price indices does not seem to me to produce a convincing case against the observation that there has been a productivity slowdown. There is no reason to suppose that if you made the same corrections on price indices before 1970 you would not have at least as much overstatement of inflation. So I think there was a productivity slowdown, I think it had an international character, it happened as much in Japan as it did in the USA, and I think that as far as anyone can tell at least half or more of it is a

mystery still. But when I say mystery I think we should distinguish between two senses of the word inexplicable (or mystery, for that matter). When I say something is inexplicable or a mystery I could mean that I cannot pin down in detail the causes of the phenomenon. But inexplicable may also mean that it is utterly shocking! How could such a thing happen? I think the productivity slowdown is inexplicable only in the first of those two senses. There is nothing in any piece of growth economics, theoretical or empirical, which says that the rate of growth of the residual is an invariant, that it cannot change from one period of time to the next. We know by the usual backward extrapolation that there cannot have been productivity growth at 1 or 2 per cent a year forever or else Oliver Cromwell would have been crawling around in skins. By the way, this goes back to a significant analytical issue. When I say that in my work in the 1950s I treated technical change as exogenous, that does not mean that I really believed at the time that it had no internal economic causes. In the very same papers I always treated population growth as exogenous, but I did know about Malthus, and there is clearly a connection between economic development and demographic patterns. What I meant by saying something is exogenous was that I do not pretend to understand this; I have nothing worthwhile to say on this so I might as well take technical change as given for reasons which are inexplicable in the first sense I mentioned before. I do not know what the determinants of technical change are in any useful detail. But technical change is not inexplicable in the second sense. I am not shocked to learn that productivity growth after 1973 is slower than before 1973, nor would I have been shocked in this sense if it had been higher.

If we take a longer-term view, going back a hundred years or so, perhaps the bigger puzzle is the above trend productivity performance of the post-war period up until the 1970s.

Exactly. I do believe that. It is a hypothesis that makes sense to me; and I can even tell a story that makes sense to explain it. But keep in mind that I am, so to speak, estimating one parameter with one degree of freedom, so there is no real test being made. The story I tell myself is as follows: from 1930 to 1947 or so, a certain amount of technological change and other improvements in productive knowledge were taking place, but could not be incorporated into the real economy, first because of the Great Depression and then because of the war. So beginning around 1950 the world had a 20-year backlog of technological improvements to incorporate into practice. After 1950 this began to happen. That seems to make perfectly good sense; but I do not believe that it is possible to test the hypothesis because there is nothing to compare it with.

Are there any strong theoretical or empirical reasons for believing that moderate inflations of less than 10 per cent have any significant adverse effects on economic growth?
I am not up on all the literature on this topic. But what I have gathered is that, at least empirically, the evidence is that rapid inflation is unconditionally bad for economic growth, but relatively slow inflation, even perhaps averaging 10 per cent annually, has no visible correlation with economic performance. I doubt that theory compels that view; but I can easily imagine that theory would be compatible with that view.

The modern phase of endogenous growth theory has now been with us for just over ten years. What do you think have been the most important developments or insights which have emerged from this research programme? Have we learned anything useful?
Less than I had hoped. My own opinion, which I think is now shared by Paul Romer, is that the early developments – the so-called *AK* models which simply amounted to saying let us assume that there are exactly constant returns to the collection of accumulatable factors of production, human and physical – all that led nowhere because it was not robust theory. It is very unlikely that growth could happen that way. If you adopt the *AK* view, it is the simplest thing in the world to say: I can show you how reducing a tax on capital will increase the growth rate, or I can show you how making leisure less attractive will increase the growth rate. But that sort of stuff went nowhere and added no real insights because it rested entirely on a linearity which is so unlikely to be true. But then when you start asking questions about what *does* govern the accumulation of technical knowledge, how could you model the accumulation of human capital, then you begin to get into really interesting issues. That is what I like about all that literature.

The current crisis involving the so-called 'Asian Tiger' economies is making big news. Their success in the past has been identified with, among other factors, export performance. What in your view is the relationship between foreign trade and growth? Has the East Asian growth 'miracle' been export-led?
Well, I am uncertain about the relationship between trade and growth. Empirically there does seem to be a relationship. I have lots of friends who have worked on this empirically and while their results differ, and some of them come up empty-handed, it appears that openness to trade favours economic performance. It is a bit less clear, at least in the literature I have read, what the source of that relationship is. The very important distinction needs to be made between factors that have growth effects and factors that have level effects. Imagine exponential growth as a linear trend on semi-log paper. You

can ask: are there forces that take a country's trend line and lift it without changing its slope, a level effect, shifting the trend roughly parallel? It is clear that anything that improves economic efficiency can do that. So trade which increases economic efficiency can almost certainly do that. If what you are looking for is something that will change the slope of the trend line, the rate of growth, then sheer efficiency gains from trade cannot do that except temporarily, not over a very long period of time. The only way you can make sense of trade having an effect on the long-term growth rate is not so much whether the country is export-led, but whether the country is in contact with the rest of the world.

So the important factor for growth is the degree of openness of an economy? Yes, openness in general and especially the will and the capacity to pick up new technology and new ideas from the rest of the world. I am absolutely clear that a positive impact from trade on efficiency and the level of output happens, and that those countries that went in for export-led growth rather than import substitution integrated themselves with the world economy and learned things. Some of it they learned from transplants, from direct foreign investment and multinationals. But whether or not there are any cases of really long-term changes in growth rates as a consequence of trade is I think very uncertain. I can easily imagine a country breaking out from being an underdeveloped stagnant economy and getting on the growth train as a result of trade. That I can easily see. But whether a country that is already growing at the same sort of rate as the OECD economies can improve its *growth rate* over a period of many decades by virtue of openness or trade, that seems to me to be unproved.

There were some interesting papers published in the first issue of the Journal of Economic Growth *[1996] by Robert Barro, Alberto Alesina and others on the relationship between democracy and economic growth, and political instability and economic growth. Barro, for example, suggests that the best way we can help poor countries is to export our economic system to them and if, as a result, their economies improve, they will tend to become more democratic. In other words economic freedom, by promoting economic growth, will eventually lead to more democratic outcomes in today's poor countries. Have you looked at this literature and developed any ideas on these issues?* That gets into questions that are too big for little old me [*laughter*]. But my reaction to that kind of literature has always been as follows. I can easily see that if you compare a democratically organized country with a country which is really tightly oligarchically organized, then the democratically organized economy is going to tap the bigger store of entrepreneurship, whereas the oligarchs in the non-democratic countries pretty soon give up any entrepre-

neurial pretensions they ever had in favour of wine, women and song or whatever [*laughter*]. There is also the question of whether you can run a modern economy in a tightly authoritarian regime or whether these are incompatible. What happens next in China is going to be the big example of this dilemma. I can understand this kind of difference. But the notion that if you could take roughly democratic countries and order them on a scale from zero to one, that going from 0.5 to 0.6 on a measure of democracy gets you a big or detectable difference in growth rate or even the level of output, that seems to be much less likely.

The greens and environmentalists are always warning everyone that the costs of economic growth will eventually outweigh the benefits. Do you ever worry about the environmental consequences of growth? Can the world sustain OECD levels of output per capita for China, South Asia, Africa and Latin America?

Yes, I worry first of all that rapid population growth will begin to encroach on the possibilities of improving productivity and on the environment. Furthermore, one of my sons is interested in these issues professionally; he likes to say that China is made of coal, that China is just one large coal deposit. Now if they were simply to burn that coal, while it might not have any effect on the growth rate of GDP as we measure GDP, it would certainly have a big effect on the growth of some rough welfare equivalent of GDP. Yes, of course I worry about these things. I worry more about that than I do about resource exhaustion, simply because we seem to be a lot further from resource exhaustion. By the way, I do think that the issue of the relationship between economic growth and the environment is, to put it crudely, probably going to boil down to a race between technology and pollution. We do not have much of a grip on the likely outcome of that issue and possibly cannot get much of a grip on it. It is foolish to be a fatuous optimist on these issues, but it is equally foolish to believe that we have come to the end of our capacity to overcome resource limitations technologically.

What sort of issues are you currently working on?

Well, not much. I am going on 74 years old and travelling a lot, as you have noticed. I do not have a long, active research agenda at the moment, although I would like to get back to one if I can. I still intend to do work in macroeconomics. The main thing that I want to work further on is what macroeconomics looks like when it takes imperfect competition seriously. Frank Hahn and I wrote an approximately unreadable book [1995] which was published a couple of years ago. There we made an attempt to outline how you might make a macro model that takes imperfect competition seriously, and possibly also takes increasing returns seriously, because increasing returns to scale are

a standard reason why competition is imperfect. We might have done reasonably well in that particular chapter, but we did not carry the model nearly far enough. In particular we did not develop it to the point where you could sensibly ask what the appropriate values are for the main parameters, if the model is to be roughly in the ball park for the US, British or German economy. I would like to go back and develop that model further. I also have a couple of ideas on growth theory, but that is another story.

PAUL M. ROMER

Copyright: Anne Knudsen

Paul Romer was born in 1955 in Denver, Colorado and obtained his BS (Maths, 1977) and his PhD (1983) from the University of Chicago. His main past posts have included: Assistant Professor at the University of Rochester, 1982–8; Professor at the University of Chicago, 1988–90; and Professor at the University of California, Berkeley, 1990–96. Since 1996 he has been Professor in the Graduate School of Business at Stanford University.

Professor Romer is best known for his influential contributions to the field of economic growth, which have led to the renaissance of economic growth analysis and, in particular, the development of endogenous growth models which highlight the importance of ideas in driving economic growth. His most widely read articles include: 'Increasing Returns and Long-Run Growth', *Journal of Political Economy* [1986]; 'Growth Based on Increasing Returns Due to Specialization', *American Economic Review* [1987]; 'Endogenous Technological Change', *Journal of Political Economy* [1990]; 'Idea Gaps and Object Gaps in Economic Development', *Journal of Monetary Economics* [1993]; 'The Origins of Endogenous Growth', *Journal of Economic Perspectives* [1994] and 'Why, Indeed, in America? Theory, History, and the Origins of Modern Economic Growth', *American Economic Review* [1996].

We interviewed Professor Romer in Chicago, in his hotel room, on 5 January 1998 while attending the annual conference of the American Economic Association.

Background Information

Where and when did you first begin to study economics?
I was a math and physics major at the University of Chicago. I took my first economics course in my senior year because I was planning to go to law school. I did well in the class and the professor encouraged me to go on to graduate school to study economics. Economics offered some of the same intellectual appeal as physics – it uses simple mathematical models to understand how the world works – and in contrast to physics, it was an area of academic study where I could actually get a job.

In some ways staying at the University of Chicago was attractive because it had a very exciting economics department but I had already been there for four years. Even though I had had very little Chicago economics training I did not think it was a good idea to spend my whole career as a student in one place, so I started in the PhD programme at MIT. There I met my wife, who was just visiting for a year from Queens University in Canada. After two years at MIT we went back to Queens to finish her final year of training in medicine. That was when I started working on growth. At the end of that year I transferred to Chicago, where my wife had a fellowship position, and completed my PhD. I finished my PhD and entered the job market in 1982 – my thesis is actually dated 1983 because it took me a year to polish it up.

As a student did you find any of your teachers to be particularly influential or inspirational?
Well, Sam Peltzman was the professor who encouraged me to switch my career path from law to economics. I shudder to think what my life would have been like if he hadn't asked to talk with me after the mid-term and I had gone on to law school. It is an episode that I try to keep in mind – that professors can be very influential, and a little bit of attention to your students as people can make a big difference in their lives.

Besides having saved me from a life in the law, Sam was also an excellent teacher. He was the first person to show me that you could take very simple tools – demand curves or indifference curves – and derive surprising insights about how the world works. Having mentioned Sam, I should also mention some other very good teachers that I had. Donald McCloskey, now Deirdre McCloskey, was the second person I had for economics. Donald, like Sam, took economics very seriously. Together, they gave me an excellent introduction to the subject. I should also mention that at Chicago, they did not offer

what is known as a 'principles' course, the watered-down, mind-numbing survey course that most universities offer as a first course in economics. At Chicago, they started right off at the intermediate microeconomics level. So I had the enormous advantage of starting off with challenging, intellectually coherent material and first-rate teachers. I was very fortunate.

Later in graduate school, when I was back at Chicago, Bob Lucas and Jose Scheinkman had a big influence on my style and the way I look at the world. They set a standard for rigour and discipline – zero tolerance for intellectual sloppiness – that I have aspired to ever since. But probably the best year of graduate school was the year I spent at Queens University because I had a lot of interaction with the faculty there. Normally as a graduate student you do not really get that much time to sit and talk with members of the faculty as colleagues. At Queens I had more of that kind of experience. Some of the people I talked with intensively during that year – Russell Davidson and James McKinnon – are terrific economists and had a big effect on my career.

Development of Macroeconomics

Are there any particular papers or books that you would identify as having a major influence on the development of macroeconomics?
For me that's too broad a question. I could list all the usual suspects, people like Keynes, and so on. I'd be more comfortable describing the contributions that have influenced my own work.

Tell us about the influences on your own research interests.
Bob Lucas brought a style to macroeconomics that had a big impact on a whole generation of people, including me. There are several papers that exemplify this style. One is his 1972 *Journal of Economic Theory* paper on 'Expectations and the Neutrality of Money'. Another would be his 1978 *Econometrica* paper 'Asset Prices in an Exchange Economy'. But his 1971 *Econometrica* paper, 'Investment under Uncertainty,' written with Edward Prescott, is probably the best example because it really brought to the forefront and crystallized for macroeconomists the connection between what we did in macroeconomics and what the rest of the profession had been doing in general equilibrium theory. In that paper Lucas and Prescott used the connection between solving optimization problems and equilibria that has become such a powerful tool in modern macroeconomics. That 1971 paper builds on the work of people like Cass [1965] and Koopmans [1965], who had been working in growth theory, and this basic approach for characterizing dynamic equilibria can be traced all the way back to Frank Ramsey's [1928] paper. Still, Lucas and Prescott took this approach much further into the core of macroeconomics. If all you have seen is the theory of investment as devel-

oped by the macro modellers and presented by the macro textbook writers, this paper is like a flash of lighting in the night that suddenly shows you where you are in a much bigger landscape.

You mentioned the influence of Bob Lucas's work. What do you think has been the lasting impact of his work, particularly the work he carried out in the 1970s for which he was awarded the Nobel Prize?
I think the deeper impact of Lucas's contributions has been on the methodology of the profession. He took general equilibrium theory and operationalized it so that macroeconomists could calculate and characterize the behaviour of the whole economy. Just as Peltzman and McCloskey took intermediate microeconomics seriously, Lucas took general equilibrium theory seriously. Many of the people doing general equilibrium theory for a living did not really seem to believe in what they were doing. They gave the impression that it was a kind of mathematical game. Economists working in trade and growth had shown us how we could use general equilibrium models, but they were not ready to bring dynamics and uncertainty into the analysis. It was economists working first in finance, then in macroeconomics, who took the theory seriously and showed economists that fully specified dynamic models with uncertainty had real implications about the world. A very important result of that methodological shift was a much greater focus on, and a much deeper understanding of, the role of expectations. But this is only part of the deeper methodological innovation. You still wouldn't know it from reading textbooks, but to research professionals, it finally is clear that you can't think about the aggregate economy using a big supply curve and a big demand curve.

One of the ironies in this revolution in thinking is that the two people who did the most to bring it about, Lucas and Robert Solow, ended up at swords' points about the substantive conclusions that this methodology had for macroeconomic policy. Solow's work has also had a huge impact on the profession, pushing us in the same direction. His work on growth also persuaded economists to take simple general equilibrium models seriously. Many people recognize the differences between Lucas and Solow over macro policy questions, but fail to appreciate the strong complementarity between their work at the methodological level. If Joan Robinson had won the day and banished the concept of a production function from professional discourse, Lucas and Prescott could never have written 'Investment under Uncertainty'.

During the 1980s, the real business cycle approach to aggregate fluctuations developed in parallel with new growth theory. How do you view that work, in particular the way it has sought to integrate the analysis of fluctuations and growth?

A lot of the progress in economics still comes from building new tools that help us understand very complicated systems. As a formal or mathematical science, economics is still very young. You might say it is still in early adolescence. Remember, at the same time that Einstein was working out the theory of general relativity in physics, economists were still talking to each other using ambiguous words and crude diagrams.

To see where real business cycle theory fits in, you have to look not just at its substance and conclusions but also at how it affected the methodological trajectory I was talking about before. You can think of a hierarchy of general equilibrium models – that is, models of the whole economy. At the top you have models of perfect competition, which are Pareto-optimal so that you can solve a maximization problem and immediately calculate the behaviour of the economy. Then, at the next level down, you have a variety of models with some kind of imperfection – external effects, taxes, nominal money, or some kind of non-convexity. In many cases you can find a way to use some of the same maximization tools to study those dynamic models even though their equilibria are not Pareto-optimal. This is what Lucas did in his 1972 paper 'Expectations and the Neutrality of Money'. Formally it is like an external effect in that model. It is also what I did in my first paper on growth.

The real business cycle guys went one step further than Lucas or I did in trying to simplify the analysis of aggregate economies. They said, 'We can go all the way with pure perfect competition and pure Pareto optimality. We can even model business cycles this way. Doing so simplifies the analysis tremendously and we can learn a lot when we do it.' My personal view – and increasingly the view of many of the people like Bob King, who worked in this area – is that at a substantive level real business cycle theory simplifies too much. It excludes too many elements that you need to understand business cycles. This doesn't mean that the initial work was bad. It just means that we are now ready to go on to the next stage and bring back in things like predetermined nominal prices. Methodologically this work helped us refine our tools so we'll do a better job of understanding predetermined prices when we bring them back into the model.

We frequently make progress in economics by seeming to take a step backwards. We assume away real problems that people have been working on in vague and confused ways, strip things down to their bare essentials, and get a better handle on the essentials using some new tools. Then we bring the complications back in. This is what Solow was doing, and what drove Robinson to distraction, when he modelled the production structure of an economy using an aggregate production function. Later we brought back many of those complications – irreversible investment, limited *ex post* substitution possibilities, and so on – back into the model. The real business cycle theorists did the same kind of thing, and during the simplification phase, they also made people mad.

Economic Growth

In Lucas's [1988] paper 'On the Mechanics of Economic Development' he comments that once you start to think about growth it is hard to think about anything else. In the introduction of their textbook Economic Growth *Robert Barro and Xavier Sala-i-Martin [1995] argue that economic growth is the part of macroeconomics that really matters. In the light of these comments by very influential macroeconomists, do you think that, on reflection, economists have in the past spent too much time trying to understand business cycles?*

That is almost right. Remember that we experienced major macroeconomic calamities in the interwar period. These depressions were sufficient to wipe out 30 to 40 years' worth of growth. Economists who grew up during this era certainly didn't have any trouble thinking about something else besides long-run growth. They naturally focused on avoiding those calamities.

So I don't think that you can make the statement that focusing on growth is more important in some absolute sense than focusing on stabilization. What I think is correct is that we now know how to avoid the kind of catastrophic events that we saw in the UK in the 1920s and in the USA in the 1930s. Those were both major mistakes in monetary policy and we now know how to avoid them. We also know how to avoid the disruptive hyperinflations of the interwar era. Recently, we have even developed better monetary rules for avoiding the less disruptive but still costly inflation of the 1970s. Once you have learnt to avoid those kinds of problems, growth stands out as the most important remaining issue on the agenda.

I do believe that there was a period in the 1960s and 1970s when macroeconomists were spending too much time looking at business cycles – the smaller cycles and fluctuations which characterized the post-war period – and too little time on growth. We should have kept working on stabilization policy, but we should also have worked on the determinants of long-run growth. Adjusting the balance is what my career has been all about.

When I teach students I try very hard to get them to get this balance right. I give them an analogy about a runner who is trying to train for a marathon. Asking whether growth is more important than stabilization is like asking whether conditioning is more important than putting on a tourniquet when the runner starts bleeding. In a sense the training and the technique of running really are what wins races. But if the runner is bleeding to death, it is pretty silly to lecture her about getting in better shape.

But now, when we look at the allocation of the profession's intellectual resources today, we are in a situation where we can learn more about how to make minor adjustments in the amplitude of cycles or in the trend rate of growth. Faced with that trade-off, it is very clear that small improvements in

the trend rate of growth can have far greater effects on the quality of life, and this area has been understudied.

Looking back, one of the reasons why economists avoided questions about growth was that our tools were not sufficiently well developed. Purely technical or mathematical issues about the existence of a solution to an infinite horizon maximization problem, transversality conditions, knife-edge behaviour and explosive growth deterred economists from asking the right kind of substantive questions about long-run growth. Now that our tools are better, we have been able to set those issues aside and make progress on the substantive questions.

The classical economists were very concerned with long-run issues such as growth. Did you find any inspiration for your work by going back and looking at the contributions of the classical economists and other early work on growth?

I did spend some time thinking about that, reading Adam Smith and Alfred Marshall. For example, I read the 1928 paper by Allyn Young, which builds on Marshall's work. I think it is in the same issue of the *Economic Journal* as Ramsey's paper. So there was a period where I spent a couple of years trying to sort out the connections between what Young and Smith were saying and what I was trying to say. I did that for a while and enjoyed it, then I stopped doing it. I am not sure I would recommend it as a research strategy for a young person, but it can be interesting and instructive.

When I started working on growth I had read almost none of the previous literature. I started very much from a clean sheet of paper and only later went back to try to figure out what other people had said. I think that in a lot of cases that is the right way to do it. If you devote too much attention to ancestor worship, you can get trapped and lose the chance to see things from a new perspective. Of course, in economics, your ancestors are still around, occupying positions of power in the profession, and they are not always happy when someone comes along and tries to take a fresh look at things.

During the whole period from the marginalist revolution in the 1870s through to the mid-1950s economists were mainly concerned with microeconomic developments and managing the birth of macroeconomics during the Great Depression. Then the issue of economic growth came back on to the scene during the 1950s. One of the puzzles is that during the period when growth theory made great advances, with the contributions of Solow in 1956 and 1957, the field of development economics seemed to evolve as an almost separate area of interest. Why did that dichotomy happen?

I am probably going to sound like a broken record here, repeating my message over and over, but the divide was methodological. The growth guys

talked math; the development guys still talked words. They diverged further and further because they could not understand each other. It was less the differences in the substantive questions they were asking than the tools they were selecting to try to address them.

Wasn't it more the case that development economists actually wanted and needed to say something about policy issues?
There was an element of that. As I said about the real business cycle theorists, sometimes you have to take a step back and simplify to make progress developing new formal tools. This is hard to do when you are in the thick of the process of trying to offer policy advice.

If you go back and read Smith, Marshall or Young, you have to be struck by what an incredibly wrenching transformation the economics profession has gone through, from operating as a purely verbal science to becoming a purely mathematical one. Remember that Allyn Young's paper came out at the same time, even in the same issue, as Frank Ramsey's. Ramsey was using tools like the calculus of variations that physicists had been using for decades. But economists were still having trouble with basic calculus. Jacob Viner needed help from his draughtsman to get the connection right between long-run and short-run average cost curves. Nowadays economists use math that is as sophisticated and as formal as the math that physicists use. So we went through a very sharp transition in a relatively short period of time. As we learned how to use mathematics we made some trade-offs. You could think of a kind of production possibility frontier, where one axis is tools and the other axis is results. When you shift effort towards the direction of building tools you are going to produce less in the way of results. So the development guys would look at Solow and say, 'What you are producing has no useful content for policy makers in the development world; you guys are just off in mathematical space wasting time while we are out here in the real world making a difference.' The tool builders should have responded by explaining the intertemporal trade-off between results and tool building and that as a result of this work we can give better policy advice in the future. Overall the right stance for the profession as a whole is one where we approve of the division of labour, where the people who specialize in those different activities can each contribute and where we do not try to force the whole profession into one branch or the other. Ideally we should keep the lines of communication open between the two branches.

Let us turn to Robert Solow's contributions. What do you see as being the main strengths and weaknesses of the Solow growth model? Some economists like Greg Mankiw [1995] would prefer to modify the Solow model rather than follow the endogenous growth path.

When it was introduced, the Solow model made several very important contributions to economics and progress in this tool-building direction. It was a very important demonstration of how you could take general equilibrium theory and apply it and say things about the real world. As I suggested before, Solow helped persuade us that there are ways to think about the equilibrium for the whole economy, using simple functional forms and simplifying assumptions, and get some important conclusions out of that. It is a very different style of general equilibrium theory from that of Arrow and Debreu and their more abstract work that was going on at the same time. Remember that Solow and Samuelson had to engage in vicious trench warfare about this time with Cambridge, England, to make the world safe for those of us who wanted to use the concept of a production function.

At the substantive level – which I think is where your question was directed – the strength of Solow's model was that he brought technology explicitly into the analysis in both his empirical paper and his theoretical paper. He had an explicit representation for technology, capital and labour. Those are the three elements that you have to think about if you want to think about growth. That was the good part. The downside was that because of the constraints imposed on him by the existing toolkit, the only way for him to talk about technology was to make it a public good. That is the real weakness of the Solow model. What endogenous growth theory is all about is that it took technology and reclassified it, not as a public good, but as a good which is subject to private control. It has at least some degree of appropriability or excludability associated with it, so that incentives matter for its production and use. But endogenous growth theory also retains the notion of non-rivalry that Solow captured. As he suggested, technology is a very different kind of good from capital and labour because it can be used over and over again, at zero marginal cost. The Solow theory was a very important first step. The natural next step beyond was to break down the public-good characterization of technology into this richer characterization – a partially excludable non-rival good. To do that you have to move away from perfect competition and that is what the recent round of growth theory has done. We needed all of the tools that were developed between the late 1950s and the 1980s to make that step.

Let me place the other strand of growth in context, the so-called *AK* versions of endogenous growth. In these models, technology is just like any other good – we might put another label on it and call it human capital or we can call it generalized capital – but technology is treated as being completely analogous with physical capital. I think that approach represented a substantive step backwards compared to the Solow model. The *AK* models are less sophisticated than the Solow model because those models do not recognize that technology is a very different kind of input. As I suggested earlier, I also

disagree with the real business cycle methodology that says 'Let us do everything with perfect competition'. Before, you could argue that there was no alternative, but that's no longer true. We have perfectly serviceable dynamic general equilibrium models with monopolistic competition and there is no reason not to use them if they capture important features of the world.

There is still a group that says 'Let's just treat technology as pure private good and preserve perfect competition'. Then there is another group of economists who, like Mankiw, say that technology is different, but we can treat it as a pure public good just as Solow did. I think that both of these positions are mistaken. There are incredibly important policy issues where the pure private-good characterization and the pure public good characterization of technology are just completely off the mark.

Wasn't your earlier work, as exemplified in your 1986 paper, more concerned with increasing returns than the determinants of technology change?
You have to look between the lines of that paper at what was going on at the methodological level, because remember, methodological and formal issues had been holding everything up. The logical sequence in my 1986 paper was to say that as soon as you think about growth, you have to think about technology. As soon as you think about technology, you have to confront the fact that there is a built-in form of increasing returns – technically, a non-convexity. Notice that is all there in Solow's model. If you look at $AF(K, L)$ you have got increasing returns in all the relevant inputs A, K and L. So up to this point, Solow and I are on the same track. You have to think of technology as a key input and one that is fundamentally different from traditional inputs. As soon as you think about that, you face increasing returns or non-convexities. Then you have to decide how to model this from a methodological point of view. Solow said treat it as a public good. There are two variants of that. One is that it comes from the sky and is just a function of time. The other is that the government could publicly provide it. I think Solow had both of those in mind and it does not really matter which you specify. What I wanted was a way to have something where there are some increasing returns but also some private provision. I wanted to capture the fact that private individuals and firms made intentional investments in the production of new technologies. So in this sense, the paper was very much about technological change. To allow for private provision, I used the concept of Marshallian external increasing returns. This lets you describe an equilibrium with price taking but still allows you to have non-convexities present in the model. That was a first provisional step. It was a way to capture the facts: there is some private control over technology, there are incentives that matter, and there are increasing returns in the background. What happened between 1986 and 1990 was that I worked hard at the mathematics of this and persuaded myself

that the external increasing returns characterization was not right either – just as the public-good assumption of Solow was not right.

Whenever you write down theories you make approximations, you take short cuts. You are always trading off the gains from simplicity against the losses in our ability to describe the world. The public-good approximation was a reasonable first step, but we needed to keep working and improve on it. The external increasing returns approximation was something of an improvement but the later monopolistic competition version [Romer, 1990] was the one that gets about the right trade-off between simplicity and relevance.

Since Solow's [1957] paper there has been a huge literature on growth accounting. What do you think have been the main substantive findings from this research?

The general progression in that area has been to attribute a smaller fraction of observed growth to the residual and a higher fraction to the accumulation of inputs. The way that literature started out was a statement that technology is extremely important because it explains the bulk of growth. Where we are now is that technology does not explain, all by itself, the majority of growth. Initially, we overstated its importance when we claimed that technological change explained 70 per cent of growth all by itself. But there are some people who would like to push this further and say there is really no need to understand technology, because it is such a small part of the contribution to growth. They argue that we can just ignore it. That is a *non sequitur*. It does not follow logically. We know from Solow, and this observation has withstood the test of time, that even if investment in capital contributes directly to growth, it is technology that causes the investment in the capital and indirectly causes all the growth. Without technological change, growth would come to a stop.

When we spoke to Bob Solow yesterday he explained why he made technology exogenous in his model. It was simply due to his lack of understanding of the causes of technological change.

That is a reasonable provisional strategy when you are dealing with a complicated world.

A great deal of attention during the past decade or so has been focused on the so-called convergence issue. At the same time as your first important endogenous growth paper was published in 1986, Moses Abramovitz and William Baumol also had papers published that drew attention to this catch-up and convergence debate. This controversy continues to draw research interest, for example, in a recent issue of the Journal of Economic Perspectives *Lant Pritchett [summer 1997] has a paper entitled 'Divergence, Big Time'.*

When we talked to Edward Prescott two days ago he was reasonably confident that convergence would eventually occur. Did this important debate influence your own thinking about growth and what are your views on this area of research?

It is very important to keep clear what the facts are. The facts are that over the time horizon that people have looked at the data, say from 1950 to the present, there is very little evidence of overall convergence. Everybody agrees about this, even if it is not always stated up front. People who describe this tendency for countries to converge are saying that if everything else were the same – if you hold all the right variables constant – then there would be a tendency for countries to converge. For example, this is one of the key results in Robert Barro's work. This is really just a refined statement of the convergence club interpretation articulated by Baumol. If you look at countries that have the same values for these variables, then they tend to converge. But it is also true that in the background, the overall progress towards reduced dispersion in per capita incomes has been very modest. Pritchett was making a useful background point. If you go back before 1950, it must be the case that there was a period where incomes diverged quite a bit – some countries moved very rapidly ahead as others were left behind. At that time, the overall distribution of income widened for a period of time. More recently, in the post-war years, the overall distribution has been roughly constant.

So why do we care about this issue? First you might care about it from a human welfare point of view, or an income distribution point of view. On those grounds there is some reason for pessimism – we really have not made that much progress in the last 30 or 40 years. You might also care about it because you think it might help you discriminate between different theories of growth – which ones are right and which are wrong. Many people have asserted that this process of conditional convergence – everything else equal, incomes converge – is consistent with a pure Solow style model, that is, one where knowledge is a public good, all technology is a public good. So they say the evidence is consistent with the public-good model of technology. That statement is correct but the evidence is also completely consistent with a model where technology is not a public good. In this interpretation, the technology gap model, flows of technology between countries are what drive the convergence process. In this explanation, the convergence you see is catching up with technology, not just catching up in the stock of capital per worker. Under the Solow model as interpreted by Mankiw and others, technology is already the same everywhere in the world. It is a public good that is in the air like a short-wave radio broadcast, so under this model there is no room for technological catch-up. It still mystifies me that people try to justify this model in the face of direct evidence about the importance of technology flows. But they certainly use the conditional convergence evidence to back up their position.

So I do not think that the convergence controversy has helped us discriminate between the different models. As a result, I think a great deal of the attention that the convergence controversy has generated has been misplaced. Prescott's assertion is that he does not think that we are going to see continued divergence. I think he is probably right about that. I personally think that these flows of technology between countries are very important forces in the big convergence episodes that we have seen. If you look at a country like Japan and ask what lies behind its very rapid convergence with the leading nations of the world, then the transfer of technology was a critical part of the process. There are grounds for optimism, looking ahead. If we can get the right institutions in place in these developing countries, the same process of flows of technologies could be unleashed and we really could see some narrowing of worldwide income inequality. If you weight it by countries the situation looks worse than if you weight it by people, at least during the last ten years. This is because the process of catching up in China will make a huge difference to the overall picture. And China is a good illustration of what is wrong with the public-good model. China had a high savings rate before the reform era. What's most different now in the sectors of manufacturing where China has been so successful has been the flow of technology into China via direct foreign investment – reforms that changed the incentives that foreign firms faced to bring technology and put it to work in China.

Did you ever look at the work of economists such as Gunnar Myrdal (1957) and Nicholas Kaldor [1970b], who tended to reject the equilibrating properties of the neoclassical model in favour of the forces of cumulative causation? In their models a lack of convergence is no surprise.

It interested me in the same way that Allyn Young interested me. I wanted to see how much there was in common between what I and what they were thinking. But it is very hard to tell, quite frankly, when you go back and read economics that is stated in purely verbal terms. There is always the danger that you read between the lines and say, oh, they had it exactly right – here is this mathematical model which shows what they were thinking. But that is usually based on a charitable reading and one that ignores some of the ambiguities and confusions. I wrote a paper like that at one point interpreting Allyn Young's paper, so one could probably do that for some of the other economists in this area. For example the big push paper by Murphy, Shleifer and Vishny [1989a] did this for some of this literature. So the right conclusion to make is that these were very smart people and they did have some good ideas, but they were working with very crude tools. I guess I would describe ancestor worship as a research strategy as probably an unproductive one [*laughter*]. But as a consumption activity it is something that can be fun.

Well, we want to keep you on the topic of ancestors for a moment. Given that your research has concentrated heavily on the influence and determinants of technological change and the importance of R&D, has the work of Joseph Schumpeter ever influenced your ideas?
No, I can honestly say that it has not. Schumpeter coined some wonderful phrases like 'creative destruction' but I did not read any of Schumpeter's work when I was creating my model. As I said, I really worked that model out from a clean sheet of paper. To be honest, the times when I have gone to try to read Schumpeter I have found it tough going. It is really hard to tell what guys like Schumpeter are talking about [*laughter*].

Too many words and not enough math?
Yes, and words are often ambiguous.

That problem has also been the source of confusion and the various conflicting interpretations of Keynes's General Theory.
Yes, right. Paul Krugman [1994c] has a nice article talking about the big push idea in development economics. When you state it now in mathematical terms, the way Murphy, Shleifer and Vishny did, you see how clearly the idea can be expressed and you wonder why someone had not done it before. I think that what it shows is that economists now are the beneficiaries of a lot of development of mathematical modes of thinking and analysis and it seems very easy to us now because we have those tools to work with. Before these techniques were available it was really very tough.

Let us go back to the issue of non-rivalry and excludability with reference to the growth of knowledge and technological change. How do you get the balance right between encouraging technological change by using incentives and yet making the new ideas and discoveries available to the rest of society? There is a trade-off problem here with respect to patent rights.
Sure. What's interesting about this question is that it is not resolved. If you take traditional private goods that are excludable and rival, we know what the best institutional arrangement is: strong property rights and anonymous markets. That's all you need. This is a remarkably important insight that economists must still communicate to the rest of the world. If people understood it, there would not be so much resistance to pricing roads, pollution or water in agriculture. Non-economists are still slow to understand how powerful the price mechanism is for allocating and producing rival goods.

But when you come to non-rival goods, we do not know what the right institutions are. It is an area that I think is very exciting because there is a lot of room for institutional innovation. One strategy is to work out a rough trade-off where you allow patent rights but you make them be narrow and

have a finite duration. You would allow partial excludability – less than full but stronger than zero excludability. We often talk as if that is the general solution. But in fact, this is not the general solution. You have to break the question down by type of non-rival good. There are some non-rival goods like the quadratic formula or pure mathematical algorithms that traditionally have been given no property rights whatsoever. There are other forms of non-rival goods like books. You will get a copyright for this book of interviews, which is a very strong form of protection. The text that you write and my words – you can take them and put a copyright on them so that nobody else can reuse them. I can not even reuse my own words without getting permission from you [*laughter*]. So that is a very strong form of intellectual property protection. What we need is a much more careful differentiation of different types of non-rival goods and an analysis of why different institutional structures and degrees of property protection are appropriate for different kinds of goods.

Patent rights or legal property rights are only a part of the story. We create other mechanisms, like subsidies for R&D. We create whole institutions like universities which are generally non-profit and government supported, that are designed to try to encourage the production of ideas. The analysis of institutions for non-rival goods is more subtle than many people realize.

For example, I have argued that it is very important to distinguish human capital from ideas – they are very different types of economic goods. Human capital is just like capital or land. It is an ordinary private good. I agree with Gary Becker on this. I think a lot of claims about human capital externalities are wrong. Nevertheless, when people conclude that we should not have any government subsidies for the production of human capital, I disagree. Why is that? It is because human capital is the crucial input into producing ideas. If you want to encourage the production of ideas, one way is to subsidize the ideas themselves. But another way is to subsidize the inputs that go into the production of ideas. In a typical form of second-best analysis, you may want to introduce an additional distortion – subsidies for scientists and engineers – to offset another – the fact that the social returns from new ideas are higher than the private returns. You create a much larger pool of scientists and engineers. This lowers the price of scientists and engineers to anybody who wants to hire their services to produce new ideas.

So in general, the optimal design of institutions is an unresolved problem. We have seen a lot of experimentation during the last 100 years. I have made the claim that the economies that will really do well in the next 100 years will be the ones that come up with the best institutions for simultaneously achieving the production of new ideas and their widespread use. I am quite confident that we will see new societal or institutional mechanisms that will get put in place for encouraging new ideas.

Research into economic growth has extended into a large number of other interesting areas in recent years. For example Alberto Alesina and Dani Rodrik [1994] have explored the relationship between inequality and growth, Robert Barro [1996], Alberto Alesina and others [1996] have explored the relationships between democracy, political stability and growth. How do you view this work? Can we help poor countries more by exporting our economic systems than our political systems, as Barro has suggested?

Let me back up a little here. One of the disciplines that formal economic theory forces on you is that you must start with an explicit conceptual framework. For example, Marshallian analysis makes us think about supply v. demand when we look at the world. General equilibrium theory forces us to split the world into preferences and the physical opportunities available to us. That split is really important and I always try to get my students to think about it when they approach a question. What do people in your model want? What are the production possibilities that are available to them?

All of growth theory has been operating under the physical opportunities question side of the model. We describe the physical opportunities as physical objects like raw materials and then start to think about ideas as recipes for rearranging these objects. When you start to think about democracy and politics, you have to start addressing the other side of the model. What is it that people want? What drives their behaviour? If you expand the concept of preferences and say that it is everything that is inside of people's heads, it includes all kinds of things that sociologists and psychologists talk about: tastes, values and norms, and so on. When you start to talk about the connections between economic growth and democracy you really have to start enquiring into these issues. Barro's assertions are based on some empirical generalizations and they are fine as far as they go, but what is missing there is any kind of theoretical understanding of the connection between economic development and political structures. This is not just a problem in economics. It is also a deep problem in political science. There are many fundamental issues that have not been addressed in political science. To begin with, why does anybody bother to vote? The standard theory that political scientists have is that people go and vote because they have a stake in the outcome and they want to influence the outcome so it goes their way – fewer taxes and more transfers, and so on. That theory contradicts itself as soon as you state it because the probability that any one voter will be decisive in an election is so trivial that the cost of going to the polls just dwarfs any possible expected gain that anyone could get from going to the polls.

So I would just assert a cautionary note here. There is a little bit of empirical evidence that suggests a connection between the level of income and democracy, but we really face an almost total theoretical vacuum in studying this question. We are unlikely to make much progress until we have

some theoretical foundations that force us to think clearly about the issues involved.

Another controversial area that has received much attention in the economic development literature is the relationship between foreign trade and growth. This is especially topical given the current crisis, which has spread throughout the 'Asian Tiger' economies that are often held up as prime examples of export-driven growth. As economists we can easily envisage an effect on the level of GDP coming from trade, but can trade influence the rate of growth?
There are two mechanisms here. From a development point of view the main thing you want to think about is this process of catching up. The key role for trade is that it lets developing countries get access to ideas that exist in the rest of the world. I tell my students that in the advanced countries of the world, we already know everything that we need to know to provide a very high standard of living for everybody in the world. It is not that we lack physical resources; it is not a lack of mass or matter that makes people in India and China poor. What makes them poor is that they do not have access to the knowledge and ideas that we have already worked out in North America, Europe and Japan for doing all the things that we do in the modern economy. The trick to make them better off is just to get that knowledge flowing into those countries. Much of it is very basic knowledge – like how to operate a distribution system so that clothes get from a factory to a store shelf so that someone can buy a shirt when they want one. How do you make sure that food does not spoil and is distributed to the right locations at the right times? How do you implement quality control systems in a manufacturing process? This is all basic knowledge but it is the stuff that raises living standards. A lot of that knowledge can be put to work in poor countries if they allow the right kinds of trade. Direct foreign investment from multinationals, in particular, is important for getting quick access to these kinds of ideas.

There is also a second issue. If you take the rich economies, OECD countries for example, the larger the market the bigger the incentives are to develop new ideas. So free trade in very large market areas creates greater incentives for innovation and therefore leads to more technological progress. If you don't think that this is true, just ask yourself how much innovation would be taking place in Silicon Valley if products made there had to be sold just in the USA, or just in California, or just in Santa Clara County? Some, to be sure, but a lot less than we see right now.

So trade matters for catching up. It also matters for sustaining growth in the leading countries.

Since growth is so important to the improvement of living standards, it is inevitable that governments will try to influence the growth rate. What should

the role of government be with respect to growth? In particular, what role do you see for monetary and fiscal policy here?
On monetary policy it is a bit like the distinction I talked about before – stopping the bleeding v. getting in shape. There is a certain amount of emergency medicine that governments have to be prepared to engage in. A lot of that amounts to an injunction to do no harm. It helps enormously if policy makers just keep from screwing up the way they did in the interwar period. But a sensible monetary policy only creates the opportunity for growth to happen; it does not make it happen. On the fiscal side, a government has to be able to pay its bills and it must keep from taxing income at such high rates that it severely distorts incentives.

There are other policies that also matter. Some of those involve creating a legal framework. What kind of institutions matter if you are in the USA? Venture capital, fluid capital markets – think of all the things that help a company like Intel come into existence and grow into a huge force. The government did not have to do anything very active but it did have to put in place structures that permitted venture capital, a new-issue stock market and so forth. Beyond that there are measures related to human capital. There is a role for government there. The modern university, as it emerged in the USA in the last century, is one that is very focused on training and practical problem solving. It is subsidized by the government. As I said before, subsidizing human capital is a very important way to indirectly subsidize technological change. So the modern university is an example of the kind of institution that the government can support.

I should add the caveat that many of the direct roles that people articulate for the government are not justified. A lot of people see endogenous growth theory as a blanket seal of approval for all of their favourite government interventions, many of which are very wrong-headed. For example, much of the discussion about infrastructure is just wrong. Infrastructure is to a very large extent a traditional physical good and should be provided in the same way that we provide other physical goods, with market incentives and strong property rights. A move towards privatization of infrastructure provision is exactly the right way to go. The government should be much less involved in infrastructure provision. So that is one area where I disagree with some of the wild-eyed interventionists. Another is the notion that the government should directly subsidize particular research programmes to produce particular kinds of ideas. If you compare that mechanism with the mechanism of subsidizing human capital and letting the market mechanism allocate where the human capital goes and what ideas get developed, the human-capital-based approach works better. Selecting a few firms and giving them money has obvious problems. How do bureaucrats get access to all the decentralized information they need if they are to decide which projects should be supported? How do

you keep rent seeking and pork barrel politics from dominating the allocation process?

A great deal of thought has been given to the design of institutions to avoid non-trivial rates of inflation. However, the relationship between inflation performance and growth performance is far from clear, especially at low rates of inflation. How do you read the evidence on this issue?
Inflation is somewhat damaging and it is probably a non-linear relationship, so the higher the rate of inflation gets the more damaging it is likely to be.

Is this due to the greater variability of inflation at higher rates?
At least partly. The variability and the higher rates both make the damage grow more than linearly. There is no trade-off, fundamentally, between growth and inflation and therefore no reason not to aim at very low levels of inflation from a growth perspective. The best place to be is at a very low level of inflation and there is no reason to accept, say, 10 per cent inflation because we think we can get some benefit in terms of long-run growth. So if you are trying to do the best job you can on growth, you basically want to aim for whatever the consensus is on minimal inflation. That will vary between zero and 2 or 3 per cent at the moment. It may not be too harmful to be up at 6 per cent instead of 2 or 3, but if it is harmful at all, why accept even that?

During the early 1970s a great deal of interest was stirred up by the book Limits to Growth *[Meadows et al., 1972]. Since then the environmental movement has become increasingly influential. Do you ever think or worry about the environmental impact of growth or the possibility of resource limitations on growth? Can the rest of the world expect to enjoy the same living standards currently enjoyed in the OECD economies without generating an environmental catastrophe?*
Environmental problems are real problems. They are cases where our current institutional structures do not put prices on physical objects that should have prices on them. When you do not have prices on fish in the sea, market incentives cause fishermen to overfish. We know that we need to institute either a price mechanism or some regulatory system that has the same effect as a price mechanism. We will face a big challenge if, for example, human sources of carbon dioxide prove to be too much for the carrying capacity of the atmosphere. We are going to have trouble implementing a worldwide price or a regulatory system to deal with this, but we will need to do it.

However, all this is very different from saying that there are long-run limits to growth. The way to think about limits is to ask, 'What does it mean to say that our standards of living are higher now or that we have more income now than we had 100 years ago?' It does not mean that we have more mass, more

pounds or kilos of material. What it means is that we took the finite resources that are available here on earth and just rearranged them in ways that made them more valuable. For example, we now take abundant silicon and we rearrange it into microchips that are much more valuable. So the question is: how much scope is there for us to take the finite amount of mass here on earth and rearrange it in ways that people will find more valuable? Here, you can make a strong case that the potential is virtually unlimited. There is absolutely no reason why we cannot have persistent growth as far into the future as you can imagine. If you implement the right institutions, the type of growth might take a slightly different form from what we anticipated. If carbon dioxide turns out to be a really big problem and we implement institutions which raise the price of carbon emissions, then cars will get smaller. Or we might drive cars somewhat less frequently, or we might rely on video conferencing, instead of driving automobiles, to meet with family and so on. We could shift to much greater reliance on renewable biomass or photovoltaics as a primary source of energy. We have the technology to do this right now. It's a more expensive way to generate electricity than burning oil and coal, but if income per capita is five to ten times higher 100 years from now, paying a bit more for energy will be a minor issue.

The bottom line is that there are pollution and other environmental problems that we will need to address. But these problems will not stop microchips from getting faster, hard disc storage densities from continuing to get higher, new pharmaceuticals from being introduced, new communications technologies from emerging, new methods for distributing goods like overnight delivery and discount retailing from emerging. All those processes will continue in the rich countries and will spread to the poor countries. In the process, the standards of living will go up for everyone.

In looking at the post-war economic growth performances of Germany and Japan compared to the UK, do you think there is anything in Mancur Olson's [1982] argument, developed in his book The Rise and Decline of Nations, *that societies which have been stable for a long time such as the UK develop organizations for collective action which are harmful to economic efficiency and dynamism?*

His conjecture is interesting, but to evaluate it we have to come back to the discussion we had earlier about production possibilities versus preferences. What Mancur tried to do was bring back into the discussion some theory about what is going on inside someone's head. He wanted to do this so he could understand the political dynamics that influence policy decisions about universities, regulations, rent seeking and so on. Those are important questions both from a development perspective and from a long-run growth perspective for advanced countries like the UK. These are important issues,

but when we think about them it is important to distinguish between assertions about the physical world and assertions about what goes on inside someone's head. Anytime you bring politics into the discussion you are crossing that divide. At that point it is always important to remind oneself that we know very little about this area. Mancur is relying on a few empirical generalizations. He looks at historical episodes where something like a revolution or a war frees things up and then you see rapid growth. He has also looked at the general process of the growth slowdown. History is never a completely reliable guide for these kinds of questions because we do not have very many observations and the current circumstances are always different from the past. I always caution someone like Mancur to be honest about the extent of our ignorance in this area, although I encourage economists to think about these questions. Just saying that the physical world presents us with enormous opportunities for growth does not mean that we will necessarily organize ourselves and take advantage of them as rapidly as we could.

Moses Abramovitz [1986], your colleague at Stanford University, has stressed the importance of what he calls 'social capability' in the catch-up process. Differences among countries' productivity levels create a potential for catch-up providing the follower countries have the appropriate institutions and technical competence. Can we operationalize a concept like social capability?
Social capability is one of those vague terms like social capital that I think would benefit from the kind of clarification that you are forced to engage in when you write down a mathematical model. It could be something that you understand in this physical opportunity side of the theoretical framework. For example, you can think of human capital as a key complementary input for technology. So just as physical capital by itself cannot explain much – neither land nor labour can themselves produce corn, but the two of them together can – it could be that human capital is the key complement for ideas or knowledge just as land is complementary to labour. Just bringing in physical capital from the rest of the world will not work if you do not have the human capital there to work with it.

You could also interpret social capability in a broader sense. You could ask whether a country has a political or social ethic or a set of norms that lets markets operate, that encourages risk taking, that supports the rule of law as opposed to either corruption or purely discretionary negotiations. You can interpret social capability in that broader sense and there are some important issues there. But when you do this, you have to recognize that you are theorizing about what goes on in someone's head.

A great deal of research and effort has been put into investigating the existence, causes and consequences of the productivity slowdown in the USA and

other advanced industrial countries. What is your personal interpretation of the findings from this research?

When I talk to students and with people from outside the university, I try to be honest about our ignorance. It is always very tempting for economists to claim more than they know. We do not know what happened with the productivity slowdown in two senses. First, I don't think we know for sure what the basic facts are. The quality of the data is such that we cannot speak with authority and answer the question about what has happened over time to the rate of growth of productivity. Second, even if there was a slowdown we do not know the reasons with any confidence. In a recent paper with Kevin Murphy and Craig Riddell [Murphy et al., 1998] I have started looking at the labour market evidence which suggests to me that technological change has proceeded at a pretty rapid but steady pace for the last three or four decades, neither slowing down nor speeding up. This calls into question some of the interpretation of the output data that we have, which does suggest that there has been a big slowdown. But all of the inferences here have to be quite tentative. You have to be realistic about what you can expect. It could be that when we get the hard numbers we will conclude that there was a productivity slowdown and we may never completely understand why it happened. I have never claimed that endogenous growth theory is necessarily going to be able to predict or explain precisely all the things that we observe. The economy is a very complicated beast and the goal for us should not be to predict within a few tenths of a percentage point the rate of growth, prospectively or retrospectively. The real test is, does the theory give us some guidance in constructing institutions that will encourage growth? Does it help us understand what kinds of things led to difference between the growth performance of the UK and the USA in the last 100 years? If the theory gives us that kind of guidance, then it has been successful and can help us design policies to improve the quality of people's lives and that is an extremely important contribution.

Where do you think the direction of research into economic growth is likely to go next or where should it go next?

I have referred a couple of times to the process of crossing the divide from thinking only about the physical opportunities to thinking about what goes on in someone's head. Once we do that more systematically, we can begin to understand the choices that individuals and societies make about growth. I believe that we already know the policies that would speed up growth in a country like India. What we need to know is why individual and collective decision procedures in India keep them from implementing these policies. This should be the next item on the research agenda.