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.

R. E. Lucas, Jr[1]

3.1 Overview

The output of economies, as measured by the gross domestic product at constant prices, tends to grow in most countries over time. Is economic growth a universal phenomenon? Why are national growth rates of the richest economies so similar? Why do some countries exhibit periods of spectacular growth, such as Japan in 1950–1973, the USA in 1820–1870, Europe after the Second World War, or China and India more recently? Why do others sometimes experience long periods of stagnation, as China did until the last two decades of the twentieth century? Do growth rates tend to converge, so that periods of above-average growth compensate for periods of below-average growth? What does this imply for levels of GDP per capita? These questions are among the most important ones in economics, for sustained growth determines the wealth and poverty of nations.

This chapter will teach us how to think systematically about growth and its determinants. It turns out that the explanation of this fundamental process is rather simple. Fig. 3.1 shows the per capita GDP of the UK since 1920. It is plotted in logarithms because, as explained in Chapter 1, the slope of the curve directly shows the growth rate. The curve reminds us of important historical events: the build-up of a war economy in the late 1930s, the massive destructions and disruptions that followed and, more recently, the excessive growth of the early 2000s that led to the 2008–12 crisis. The figure also plots the trend. The amazing observation is that, in the end of it all, the economy always comes back to its trend. No matter how profound are these historical episodes, Britain has been growing at 2.1% per annum. This suggests that there exist some fundamental forces that, if left free to operate, keep an economy increasing indefinitely.

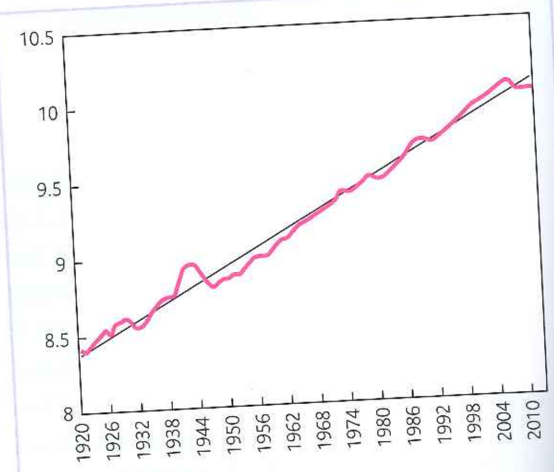


Fig. 3.1 GDP per capita in the UK 1920–2011

Using a logarithmic scale, we show the evolution of per capita GDP along with a trend that corresponds to a 2.1% growth rate. We see important events: the Great Depression in the early 1930s, the build-up of a war economy at the end of the same decade, the sharp fall as a result of the Second World War and the overheated 'bubble economy' of the 2000s. Remarkably, the GDP returns over time to its seemingly immutable growth trend.

Source: Maddison (1995), AMECO

Three main forces can be identified. First, investment can add to the capital stock, and a greater capital stock enables workers to produce more. Second, the working population or labour force can grow, which means that more workers are potentially available for market production. This growth can arise for many reasons—increases in births two or three decades ago, immigration now, or increased labour force participation by people of all ages, especially by women. The third reason is **technological progress**. As knowledge accumulates and techniques improve, workers and the machines they work with become more productive. For both theoretical and empirical reasons, technological progress turns out to be the ultimate driver of economic growth. Because it is such an important topic, a detailed discussion of technological progress will be postponed to Chapter 4.

[1] Robert E. Lucas, Jr (1937–), Chicago economist and Nobel Prize Laureate in 1995, is generally regarded as one of the most influential contemporary macroeconomists. Among his many fundamental contributions to the field, he has researched extensively the determinants of economic growth.

3.2 Thinking about Economic Growth: Facts and Stylized Facts

3.2.1 The Economic Growth Phenomenon

Despite setbacks arising from wars, natural disasters, or epidemics, economic growth seems like an immutable economic law of nature. Over the centuries, it has been responsible for significant, long-run material improvements in the way the world lives. Table 3.1 displays the annual rate of increase in real GDP—the standard measure of economic output of a geographic entity—for various periods in a number of currently

Table 3.1 The Growth Phenomenon

	Average rates of growth in GDP (% per annum)							Av. Growth GDP per capita 1820–2010 (% per annum)
	1820–1870	1870–1913	1913–50	1950–73	1973–2001	2001–2010	1820–2010	
Austria	1.4	2.4	0.2	5.2	2.5	2.2	2.0	1.6
Belgium	2.2	2.0	1.0	4.0	2.1	2.0	2.1	1.5
Denmark	1.9	2.6	2.5	3.7	2.0	1.7	2.4	1.5
Finland	1.6	2.7	2.7	4.8	2.4	2.3	2.6	1.8
France	1.4	1.6	1.1	4.9	2.3	1.9	1.9	1.6
Germany	2.0	2.8	0.3	5.5	1.8	1.7	2.2	1.6
Italy	1.2	1.9	1.5	5.5	2.3	1.7	2.1	1.5
Netherlands	1.7	2.1	2.4	4.6	2.4	2.2	2.4	1.4
Norway	2.2	2.2	2.9	4.0	3.4	2.9	2.7	1.9
Sweden	1.6	2.1	2.7	3.7	1.9	1.9	2.3	1.6
Switzerland	1.9	2.5	2.6	4.4	1.2	1.3	2.4	1.7
United Kingdom	2.0	1.9	1.2	2.9	2.1	2.0	1.9	1.4
Japan	0.4	2.4	2.2	8.9	2.7	2.2	2.6	1.8
United States	4.1	3.9	2.8	3.9	3.0	2.7	3.5	1.7

Source: Maddison (2007) 'Historical Statistics for the World Economy: 1–2003 AD'.

Box 3.1 China and the Chinese Puzzle of Economic Growth

At the end of the fourteenth century, China probably was the world's most advanced economy. While Europe was just beginning to recover from centuries of inward-looking backwardness and relative decline, Chinese society had reached a high degree of administrative, scientific, and economic sophistication. Innovations such as accounting, gunpowder, the maritime compass, moveable type, and porcelain manufacture are just a few attributable to the Middle Empire. Marco Polo was one of many famous European traders who tried to break into the Chinese market. According to crude estimates by economic historian Angus Maddison, fourteenth-century Western Europe and China were on roughly equal footing in terms of market output—and many experts claim the Chinese were technically more advanced.[2] Yet over the next six centuries, standards of living increased 25-fold in Western Europe compared with only sevenfold in China.

Most of that sevenfold increase in GDP per capita has occurred in China over the last three decades. After adopting far-reaching market economy reforms in the 1980s, economic growth has averaged a phenomenal 10% per annum since 1990. At this rate, the economy doubles in size every seven years. If this growth continues, China will easily reach the standard of living of poorer EU countries by 2025.

The Chinese growth phenomenon raises a host of intriguing questions. Why did China stagnate for centuries, while Europe flourished? Why did China literally explode in the 1990s? It is widely agreed that the Chinese success story would have been impossible without China's shift from isolationism to a policy of openness to international trade and foreign direct investment. Almost as a converse proposition, some historians associate the economic stagnation of China after the fifteenth century with the grounding of 3,500 great sailing ships of the Ming dynasty in 1433, the world's largest naval expeditionary fleet under the command of Admiral Zheng He. A policy of 'inward perfection', fear of Mongol threats, lack of government funding, and a deep mistrust of merchant classes which benefited most from the international excursions of the Imperial 'Treasure Fleet', all led China to close itself off from foreign influences, with disastrous consequences. For many economists, this is a warning shot about potential risks of an unquestioning anti-globalization movement. In Chapter 4, we revisit the theme of international trade and economic growth in more detail.

wealthy countries since 1820. (The early data are clearly rough estimates.) Over almost two centuries, GDP has increased by 60- to 100-fold or more, while per capita GDP has increased by 12- to 30-fold. Our grandparents are right when they say that we are much better off than they were.

The table also reveals that the growth process is not necessarily smooth. We will see that this variation reflects the effect of wars, colonial expansion and annexation, and dramatic changes in population as well as political, cultural, and scientific revolutions. It also reflects the business cycle. Despite these swings, it is striking that the overall average growth of GDP per capita is remarkably similar across these countries and over time. Indeed, the last column shows that *per capita* growth since 1820 has been 1.5–1.9% per annum.

Small average annual changes displayed in Table 3.1 cumulate surprisingly fast. The advanced economies of the world grow by roughly 2–4% per year. A growth rate difference of 2% per annum increases a gap by 49% in 20 years, and by 170% in half a century. The recent phenomenal growth successes of China and India and the troubling slowdowns in Western Europe and Japan show that growth is by no means an automatic birthright. Moreover, fortunes can change: Box 3.1 reminds us that China was a leading world economy in the fourteenth century, only to fall into half a millennium of decline and stagnation.

3.2.2 The Sources of Growth: The Aggregate Production Function

Growth theory asks how sustained economic growth across nations and over time is possible. Do we produce more because we employ more inputs, or because the inputs themselves become more productive over time, or both? What is the contribution of each factor? The most important tool we will use to think about these questions is the **production function**. The production function relates the output of an economy—its GDP—to productive inputs. The two most important productive inputs are the physical capital stock (equipment and structures), represented by K, and labour employed, represented by L. The capital stock includes factories, buildings, and machinery as well as roads and railroads, electricity, and telephone networks. Employment or labour is the total number of hours worked in a given period of time. The labour measure L is the product of the average number of workers employed (N) during a period (usually a year) and the average hours (h) that they work during that period ($L = Nh$). We speak of person-hours of labour input.[3] Symbolically, the production function is written:

$$Y = F(\underset{+}{K}, \underset{+}{L}). \tag{3.1}$$

The plus (' + ') signs beneath the two inputs signify that output rises with either more capital or more labour.[5]

The production function is a useful, powerful, and widely used short-cut. It reduces many and complex types of physical capital and labour input to two. In microeconomics, the production function helps economists study the output of individual firms. In macroeconomics, it is used to think about the output of an entire economy. Box 3.2 presents and discusses

2 Maddison (1991: 10).

Box 3.2 For the Mathematically Minded: The Cobb–Douglas Production Function

The use of mathematics in economics can bring clarity and precision to the discussion of economic relationships. An illustration of this is the notion of a production function, which formalizes the relationship between inputs (capital and labour) and output (GDP). One particularly well-known and widely-used example is the Cobb–Douglas production function:

$$Y = K^{\alpha}L^{1-\alpha}, \tag{3.2}$$

where α is a parameter which lies between 0 and 1, and is called the elasticity of output with respect to capital: a 1% increase in the capital input results in an α increase in output.[4] Similarly $1 - \alpha$ is the elasticity of output with respect to labour input. It is easy to see that the Cobb–Douglas production function possesses all the properties described in the text.

Diminishing marginal productivity

The marginal productivity of capital is given by the derivative of output with respect to capital K: $\partial Y/\partial K = \alpha K^{\alpha-1}L^{1-\alpha} = \alpha(L/K)^{1-\alpha}$. Since $\alpha < 1$, the marginal product of capital is a decreasing function of K and an increasing function of L. Similarly, the marginal productivity of labour is given by $\partial Y/\partial L = (1-\alpha)(K/L)^{\alpha}$, which is increasing in K and decreasing in L.

Constant returns to scale

The Cobb–Douglas function has constant returns to scale: for a positive number t, which can be thought of as a scaling factor,

$$(tK)^{\alpha}(tL)^{1-\alpha} = t^{\alpha}t^{1-\alpha}K^{\alpha}L^{1-\alpha} = tK^{\alpha}L^{1-\alpha} = tY. \tag{3.3}$$

Intensive form

The intensive form of the Cobb–Douglas production function is obtained by dividing both sides of (3.2) by L, which is the same as setting $t = 1/L$ in equation (3.3), to obtain:

$$Y/L = y = (K^{\alpha}L^{1-\alpha})/L = K^{\alpha}L^{-\alpha} = (K/L)^{\alpha} = k^{\alpha}, \tag{3.4}$$

where $k = K/L$ and $y = Y/L$ are the intensive form measures of input and output defined in the text. Since $\alpha < 1$, the intensive form production is indeed well-represented by Figure 3.2.

3 Since output and labour inputs are flows, they could also be measured per quarter or per month, but should be measured over the same time interval. Note that capital is a stock, usually measured at the beginning of the current or end of the last period. We discussed the important distinction between stocks and flows in Chapter 1.

4 To see this, note that the elasticity of output with respect to capital is defined as $(dY/dK)(K/Y)$ and is given by $(\alpha K^{\alpha-1}L^{1-\alpha})(K^{1-\alpha}L^{\alpha-1}) = \alpha$. Similarly, $1 - \alpha$ is the elasticity of output with respect to the labour input.

5 Formally, the first partial derivatives $FK(K, L) \equiv \partial F(K, L)/\partial K$ and $FL(K, L) \equiv \partial F(K, L)/\partial L$ are positive.

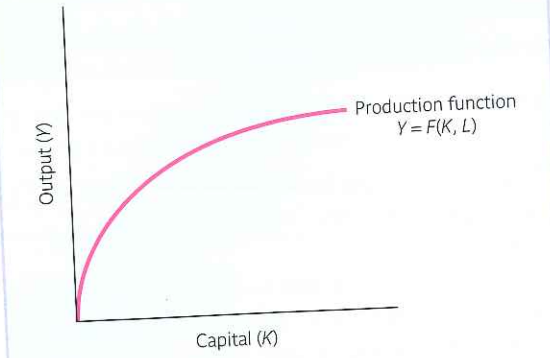


Fig. 3.2 The Production function

Holding labour input L (the number of hours worked) unchanged, adding to the capital stock K (available productive equipment) allows an economy to produce more, but in smaller and smaller increments.

the characteristics of a widely-used production function, the Cobb–Douglas production function.

The production function is a *technological relationship*. It does not reflect the profitability of production, and it has nothing to do with the quality of life or the desirability of work. It is meant to capture the fact that goods and services are produced using factors of production: here, equipment and hours of labour. In the following, we describe some basic properties that are typically assumed for production functions.

Marginal productivity

Consider an economy producing output with workers and a stock of capital equipment. Then imagine that a new unit of capital—a new machine—is added to the capital stock, raising it by the amount ΔK, while holding labour input constant.[6] Output will also rise, by ΔY. The ratio $\Delta Y/\Delta K$, the amount of new output per unit of incremental capital, is called the economy's **marginal productivity**. Now imagine repeating the experiment, adding capital again and again to the existing stock, always holding labour input constant. Should we expect output to increase by the same amount for each additional increment of capital?

Generally, the answer is no. As more and more capital is brought into the production process and the number of hours worked remains the same, each new piece of equipment is allocated less of workers' time. As one might expect, the increases in output become smaller and smaller. This is called the principle of **diminishing marginal productivity**.[7] It is represented in Figure 3.2, which shows how output rises with capital, holding labour unchanged. The flattening of the curve illustrates the assumption. In fact, the slope of the curve is equal to the economy's marginal productivity.

It turns out that the principle of diminishing marginal productivity also applies to the labour input. Increasing the employment of person-hours will raise output; but output from additional person-hours declines as more and more labour is being applied to a fixed stock of capital.

Returns to scale

Output increases when either inputs of capital or labour increases. But what happens if both capital and labour increase in the same proportion? Suppose, for example, that the inputs of capital and labour were both doubled—increased by 100%. If output doubles as a result, the production function is said to have **constant returns to scale**. If a doubling of inputs leads to more than a doubling of output, we observe **increasing returns to scale**. **Decreasing returns** is the case when output increases by less than 100%. It is believed that decreasing returns to scale are unlikely. Increasing returns, in contrast, cannot be ruled out, but we will ignore this possibility until Chapter 4. In fact, the bulk of the evidence points in the direction of constant returns to scale.

With constant returns we can think of the link between inputs and output—the production function—as a zoom lens: as long as we scale up the inputs, so does the output. In this case, an attractive property of constant returns production functions emerges: output per hour of work—the **output-labour ratio** (Y/L)—depends only on capital per hour of work—the **capital-labour ratio** (K/L). This simplification allows us to write the production function in the following intensive form:[8]

$$y = f(k), \tag{3.5}$$

where $y = Y/L$ and $k = K/L$. The output–labour ratio Y/L is also called the average productivity of labour: it says how much, on average, is being produced with one unit (one hour) of work.[9] The capital–labour ratio K/L measures the capital intensity of production.

The intensive-form production function is depicted in Figure 3.3. Because of diminishing marginal productivity, the curve becomes flatter as the capital–labour ratio increases. The intensive-form representation of the production function is convenient because it expresses the average productivity of labour in an economy as a function of the average stock of capital with which that labour is employed. In a world of constant returns, the intensive form production function says that standards of living do not depend on the absolute size of an economy. Indeed, Ireland, Singapore, and Switzerland have matched or exceeded the per capita GDP of the USA, the UK, or Germany.

3.2.3 Kaldor's Five Stylized Facts of Economic Growth

At this point it will prove helpful to look at the data: How have inputs and outputs in real-world economies changed over time? In 1961, the British economist Nicholas Kaldor (1908–1986) studied economic growth in many countries over long periods of time and isolated several **stylized facts** about economic growth which remain valid to this day. Stylized facts are empirical regularities found in the data. Kaldor's stylized facts will organize our discussion of economic growth and restrict our attention to theories which help us to think about it, just as a police detective uses clues to limit the number of possible suspects in a criminal investigation.

The first of Kaldor's stylized facts concerns the behaviour of output per person-hour and capital per person-hour.

Stylized Fact No. 1: both output per capita and capital intensity keep increasing

The most remarkable aspect of the growth phenomenon is that real GDP seems to grow without bound. Yet labour input, measured in person-hours of work (L), grows much more slowly than both capital (K) and output (Y). Put differently, both average productivity (Y/L) and capital intensity (K/L) keep rising. Because income per capita (Y/N) is closely related to average productivity or output per hour of work (Y/L), economic growth implies a continuing increase in material standards of living.[10] Figure 3.4 presents the

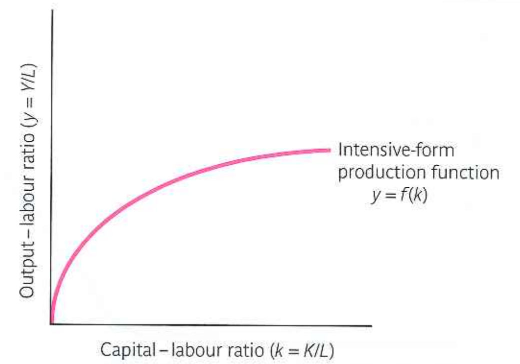


Fig. 3.3 The Production Function in Intensive Form

The production function shows that the output–labour ratio y grows with the capital–labour ratio k. Its slope is the marginal productivity of labour since with constant returns to scale $\Delta Y/\Delta K = \Delta y/\Delta k$. The principle of declining marginal productivity implies that the curve becomes flatter as k increases.

6 Throughout this book, the symbol 'Δ' (Greek delta) is used to denote a step change in a variable.

7 Following footnote 5, this feature of the production function implies mathematically that the second partial derivatives $F_{KK}(K, L) \equiv \partial^2 F(K, L)/\partial K^2$ and $F_{LL}(K, L) \equiv \partial F^2(K, L)/\partial L^2$ are negative.

8 The constant returns property implies that if we scale up K and L by a factor t, Y is scaled up by the exactly same factor—for all positive numbers t, it is true that $tY = F(tK, tL)$. In the text we use the case $t = 2$; we double all inputs and produce twice as much. If we choose $t = 1/L$, we have $y = Y/L = F(k, 1)$. Rename this $f(k)$ because $F(k, 1)$ depends on k only. The intensive production function $f(k)$ expresses output produced per unit of labour (y) as a function of the capital intensity of production (k).

9 It is important to recall the distinction between average productivity (Y/L) and marginal productivity ($\Delta Y/\Delta L$).

10 If population size N and total hours worked L change in the same proportion, the growth rates of Y/N and Y/L are identical. In developed (mature) economies, people tend to work less hours and less years, so Y/L grows more slowly than Y/N. Formally, $\Delta(Y/L)/(Y/L) = \Delta Y/Y - \Delta L/L$ and $\Delta(Y/N)/(Y/N) = \Delta Y/Y - \Delta N/N$, so the difference between the growth rates of per capita output Y/N and average productivy Y/L is $\Delta N/N - \Delta L/L$. Section 3.6.3 examines this difference.

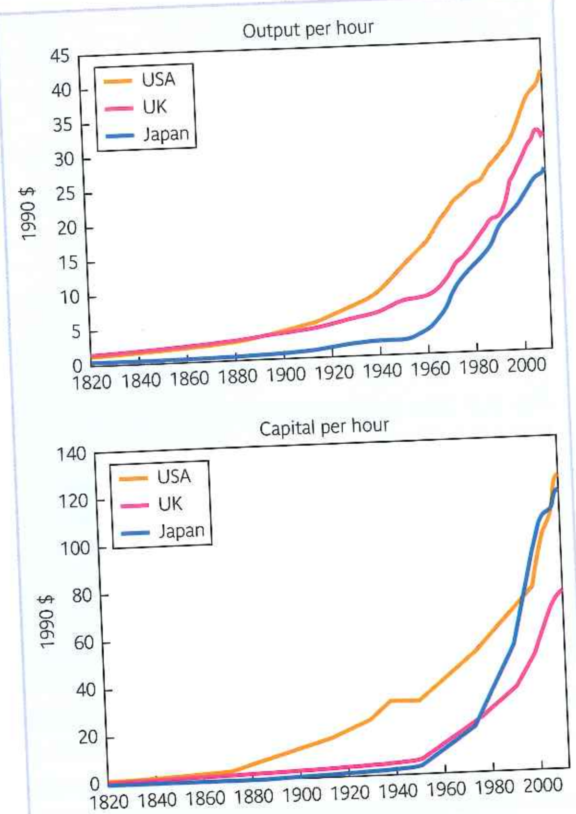


Fig. 3.4 The Output–Labour and Capital–Labour Ratios in Three Countries

Output–labour and capital–labour ratios are continuously increasing. Growth accelerated in the USA in the early twentieth century, and after 1950 in Japan and the UK.

Sources: Maddison (1995); Groningen Total Economy Database, (www.ggdc.net), OECD, *Economic Outlook*, chained.

Table 3.2 Capital–Output Ratios (K/Y), 1913–2009

	1913	1950	1973	1992	2009
France	1.6	1.7	1.8	2.3	2.6
Germany	2.3	2.1	2.4	2.3	2.6
Japan	1.0	1.8	1.7	3.0	3.9
UK	1.0	1.1	1.7	1.8	1.9
USA	2.9	2.3	2.1	2.4	2.3

Source: Maddison (1995); 2009 data from *Economic Outlook*, OECD.

evolution of the output–labour and capital–labour ratios in three important industrial economies.

Stylized Fact No. 2: the capital–output ratio exhibits little or no trend over time

As they grow in a seemingly unbounded fashion, the capital stock and output tend to track each other. As a consequence, the ratio of capital to output (K/Y) shows little or no systematic trend. This is apparent from Figure 3.4, but Table 3.2 shows that it is only approximately true. For example, while output per hour in the USA has increased by roughly 600% since 1913, the ratio of capital to output actually fell slightly over the same period. At any rate, the capital–output ratio may not be exactly constant, but it is far from exhibiting the steady, unrelenting increases in average productivity and capital intensity described in Stylized Fact No. 1.

Stylized Fact No. 3: hourly wages keep rising

The long-run increases in the ratios of output and capital to labour (Y/L and K/L) mean that, over time, an hour of work produces ever more output. Simply put, workers become more productive. It stands to reason that their wages per hour should also rise (the theoretical reasoning will be shown more formally in Chapter 5), and this is indeed the case. Growth evidently delivers ever-increasing living standards for workers.

Stylized Fact No. 4: the rate of profit is trendless

Note that the capital–output ratio (K/Y) is just the inverse of the average productivity of capital (Y/K). The absence of a clear trend for the capital–output ratio (K/Y) implies that average capital productivity too is trendless: over time, the same amount of equipment delivers about the same amount of output. It is to be expected therefore that the rate of profit does not exhibit a trend either. This stands in sharp contrast with labour productivity, whose secular increase allows a continuing rise in real wages. Yet income flowing to owners of capital has increased, but only because the stock of capital itself has increased. Indeed, with a stable rate of profit, income from capital increases proportionately to the capital stock.

Stylized Fact No. 5: the relative income shares of GDP paid to labour and capital are trendless

We just saw that incomes from labour and capital increase secularly. Surprisingly perhaps, it turns out that they also tend to increase at about the same rate, so that the distribution of total income (GDP) between capital and labour has been relatively stable. In other words, the labour and capital shares have no long-run trend. We will have to work a little harder to explain this remarkable fact.

3.2.4 The Steady State

Stylized facts are not meant to be literally true at all times, certainly not from one year to the other. Instead, they highlight central tendencies in the data. As we study growth, we are tracking moving targets, variables that keep increasing all the time, apparently without any upper limits. Thinking about moving targets is easier if we can identify stable relationships among them. This is why Kaldor's stylized facts will prove helpful. Another example of this approach is given by the evolution of GDP: it seems to be growing without bounds, but could its growth rate be roughly constant? The answer is yes, but only on average, over five or ten years or more. In Chapter 1, we noted the important phenomenon of business cycles, periods of fast growth followed by periods of slow growth or even declining output. As we look at secular economic growth, we are not interested in business cycles. We ignore shorter-term fluctuations—compare Figures 1.5 and 1.6 in Chapter 1—and focus on the long run.

This is why it is useful to try and imagine how things would look if there were no business cycles at all. Such a situation is called a **steady state**. Its central feature is that certain variables, such as GDP growth rates, or ratios of key variables, such as the capital–output ratio or the labour income share—are constant. Just as the stylized facts are not to be taken literally, think of the steady state as the long-run average behaviour that we never reach, but fluctuate about. From the perspective of 10 years ago, we thought of today as the long run, but now we can see all the details that were unknown back then. Given that modern GDPs double every 10–30 years, a temporary boom or recession which shifts today's GDP by one or two percentage points amounts to little in the greater order of things, the powerful phenomenon of continuous long-run growth. Steady states—and stylized facts—are not just convenient ways of making our lives simpler; they are essential tools for distinguishing the forest from the trees.

3.3 Capital Accumulation and Economic Growth[11]

3.3.1 Savings, Investment, and Capital Accumulation

Kaldor's first stylized fact highlights a relationship between output per hour and capital per hour. This link is in fact predicted by the production function in its intensive form. It suggests that a good place to start if we want to explain economic growth is to understand why and how the capital stock rises over time. We will thus first study how the savings by households—foregone consumption—is transformed into investment in capital goods, which causes the capital stock to grow.

The central insight is delivered by the circular flow diagram presented in Chapter 2 (Figure 2.3). GDP represents income to households, either to workers as wages and salaries, or to owners of firms as distributed profits. Households and firms save part of their income. These savings flow into the financial system—banks, stock markets, pension funds, etc. The financial system channels these resources to borrowers: firms, households, and the government. In particular, firms borrow—including from their own savings—to purchase capital goods used in production. This expansion of productive capacity, in turn, raises output, which then raises future savings and investment, and so on. This is the machinery behind the growth phenomenon.

To keep things simple, let us first assume that the size of the population, the labour force, and the numbers of hours worked are all constant. At this stage, we ask some fundamental questions: can

[11] This section presents the **Solow growth model**, in reference to Nobel Prize Laureate Robert Solow of the Massachusetts Institute of Technology.

capital accumulation proceed without bound? Does more saving mean faster growth? And since saving means postponing consumption, is it always a good idea to save more?

3.3.2 Capital Accumulation and Depreciation

Let us start from the national accounts of Chapter 2. Identity (2.7) shows that investment (I) can be financed either by private savings by firms or households (S), by government savings (the consolidated budget surplus, or $T - G$), or the net savings of foreigners (the current account deficit, $Z - X$):

$$I = S + (T - G) + (Z - X). \qquad (3.6)$$

As a description of the long-run or a steady state, suppose that the government budget is in balance ($T = G$), and the current account surplus equals zero ($Z = X$). In this case, the economy's capital stock is ultimately financed by savings of resident households.[12] More precisely, we reach the conclusion that, in the steady state, $I = S$. Investment expenditures are financed entirely by domestic savings. This is a first explanation of the growth phenomenon: we save, we invest, we grow. As a first approximation, let s be the fraction of GDP which households save to finance investment. Equality of investment and saving implies:

$$I = sY \text{ and therefore } I/L = sY/L = sy = sf(k). \qquad (3.7)$$

This relationship is shown in Figure 3.5 as the **saving schedule**. It expresses national savings as a function of national output and income. The saving schedule lies below the production function because saving is just a fraction of GDP.

We next distinguish between **gross investment**, the amount of money spent on new capital, and **net investment**, the increase in the capital stock. Gross investment represents new additions to the physical capital stock, but it does not represent the net change of the capital stock because, over time, previously installed equipment **depreciates**—it wears out, loses

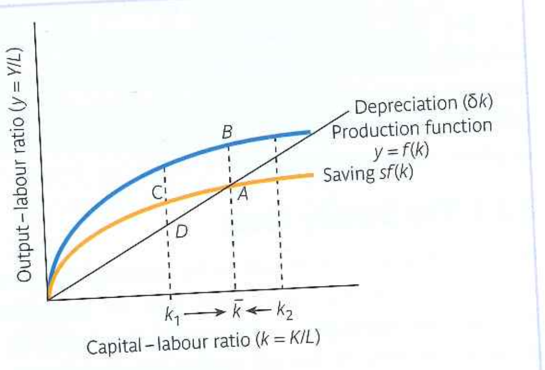


Fig. 3.5 The Steady State

The capital – labour ratio stops changing when investment is equal to depreciation. This occurs at point A, the intersection between the saving schedule $sf(k)$ and the depreciation line δk. The corresponding output – labour ratio is determined by the production function $f(k)$ at point B. When away from point A, the economy moves towards its steady state. Starting below the steady state at k_1, investment (point C) exceeds depreciation (point D) and the capital – output ratio will increase until it reaches its steady-state level $\bar{k}$.

some of its economic value, or becomes obsolescent. The fraction of the capital stock that is routinely lost each period is called **depreciation**, and the proportion of the total stock lost each period δ is called the **depreciation rate**. The depreciation rate for the overall economy is fairly stable and will be taken as constant. Logically, this implies that the more capital is in place, the more depreciation will occur. Because it is proportional to the capital stock, depreciation is represented in Figure 3.5 by a ray from the origin, the **depreciation line**, with slope δ.

Whenever gross investment exceeds depreciation, net investment is positive and the capital stock rises. If gross investment is less than depreciation, the capital stock falls. While it may seem odd to imagine a shrinking capital stock, it is not uncommon to observe in declining industries or regions. Net investment is therefore:

$$\Delta K = sY - \delta K \qquad (3.8)$$

or equivalently, written in intensive form:

$$\Delta k = sy - \delta k.$$

12 This need not be true for a region within a nation: the capital stock of southern Italy, eastern Germany, or Northern Ireland may well be financed by residents of other parts of their countries. Yet even these financing imbalances are unlikely to be sustainable for the indefinite future.

We see that the net accumulation of capital per unit of labour is positively related to the savings rate s and negatively related to the depreciation rate δ. How about the link between capital accumulation Δk and capital intensity k? It is ambiguous. On the one hand, a larger capital stock means a higher income level ($y = f(k)$), and therefore more savings and more gross investment. On the other hand, more capital means more depreciation and less net investment. This ambiguity is a central issue in the study of economic growth and will be addressed in the following sections.

3.3.3 Characterizing the Steady State

Let us summarize what we have done up to now. The production function (3.5) relates an economy's output to inputs of capital and labour. Its intensive form, presented in Figure 3.4 and Figure 3.5, relates the output–labour ratio to the capital–labour ratio. According to equation (3.8), capital accumulation is also driven by the output–labour ratio. Putting all these pieces together, we find that capital accumulation (Δk) is determined by stock of capital already accumulated up to that point (k):

$$\Delta k = sf(k) - \delta k. \qquad (3.9)$$

In Figure 3.5, Δk is the vertical distance between the savings schedule $sf(k)$ and the depreciation line δk. It represents the net change in the capital stock per unit of labour input in the economy. The sign of Δk tells us where the economy is heading. When $\Delta k > 0$, the capital stock per capita is rising and the economy is growing, since more output can be produced. When $\Delta k < 0$, the capital stock per capita and output per capita are both declining. At the intersection of the saving schedule and the depreciation line (point A), gross investment and depreciation are equal, so the capital–labour ratio (point B) no longer changes. The capital stock is thus similar to the level of water in a bathtub when the drain is slightly open: gross investment is like the water running through the tap, while depreciation represents the loss of water through the drain. The newly accumulated capital exactly compensates that lost to depreciation—the water flows into the bathtub at the same speed as it leaks out. This is the steady state.

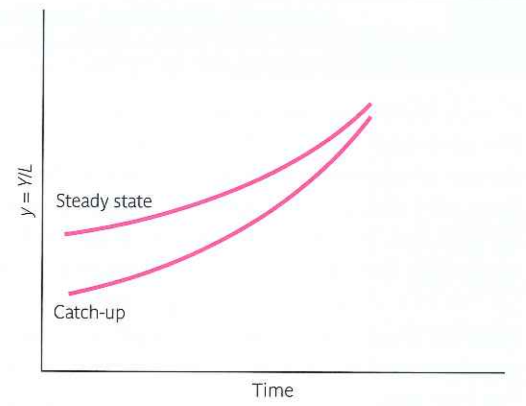


Fig. 3.6 Catching up

Catching up occurs when a country starts below its steady state GDP (lower curve) and embarks on a faster growth path, approaching the steady state (the upper curve) from below.

Wherever it starts, the economy will spontaneously gravitate to the steady state and stay there. Suppose, for instance, that the economy is to the left of the steady-state capital–labour ratio $\bar{k}$, say at the level k_1.[13] Figure 3.5 shows that gross investment $sf(k_1)$ at point C exceeds depreciation δk_1 at point D. According to (3.9), the distance CD represents net investment, the increase in the capital–labour ratio k, which rises towards its steady-state level $\bar{k}$.

Can the capital stock proceed beyond $\bar{k}$, going all way to say, k_2? It turns out that it cannot. As the economy gets closer to point A, net investment becomes smaller and smaller as the steady state is approached; it equals nil precisely when $k = \bar{k}$. This process, called catch-up, is represented in Figure 3.6 and illustrated in Box 3.3. To see how the economy behaves when capital is above its steady state, consider $k_2 > \bar{k}$. Gross investment $sf(k_2)$ is less than depreciation δk_2, the capital–labour ratio declines, and we move leftward towards $\bar{k}$, the economy's stable resting point. Later we shall see that the result that capital per capita and output per capita eventually stabilizes carries over when we account for population growth.

13 In general, steady-state values of variables will be indicated here with an upper bar, e.g. $\bar{k}$, $\bar{y}$, etc.

Box 3.3 Growth Miracles Eventually Come to an End

In each period, we see 'growth miracles', countries that go through a period of rapid growth. The USA in the late nineteenth century and Australia in the early twentieth century went through this catch-up phase. For historical reasons, a country with a low capital–labour ratio unleashes the process described in Figure 3.6. The USA and Australia were initially poor colonies. Table 3.3 show some more recent examples. Europe recovered in the 1960s from wartime destructions. The next wave included Japan, also recovering from the war and economic backwardness. Then came China, a very poor country after centuries of economic and political repression. Catch-up started once it adopted fundamental economic reforms and opened to international trade in the late 1970s under the leadership of Deng Xiaoping. More recently, India and Vietnam have joined in. Much of Latin America has also started to catch up over the last decade or two. Africa, finally, seems about to extricate itself from poverty as well.

These are amazing episodes. A 10% annual growth rate means that standards of living double every seven years. Why do these growth explosions take off, and how long do they last? Growth theory tells us that catch-up occurs spontaneously when people save and firms borrow these savings to invest and raise the capital stock, and when firms can adopt state-of-the-art technology. This requires adequate financial systems and economic openness. Importantly, it requires a reliable system of property rights, which in turn requires effective legal systems and the absence of armed conflict. As for the latter question, Figure 3.6 confirms the answer given by theory: until the economy reaches the steady state. This happened to Europe in the 1970s, to Japan a decade later. China is still poor and should keep growing for a decade or two, albeit at a decreasing rate.

3.3.4 The Role of Savings for Growth

We now show that the more a country saves, the more it invests; the more it invests, the higher is its steady-state capital–output ratio; and the larger its capital–output ratio, the higher its output–labour ratio will be in the steady state. This long-run proposition implies that countries that save and invest a lot have high per capita incomes. Is this true? Figure 3.7(*a*) looks at investment rates and GDP per capita for the whole world and is suggestive of such a relationship. For example, the poor countries of Africa typically invest little, in contrast to the richer countries of Europe and Asia. Yet, the link is hardly strong. In addition, Figure 3.7(*b*) shows that the investment rate fails to account for differences in economic growth between countries. Obviously, our story is too simple and we will soon put more flesh on the bare bones that we have just assembled.

Still, at this stage, we reach an important result: savings and investment are more closely linked to steady-state *levels* of output rather than their growth rates. This means that nations which save more should have higher standards of living in the steady state, but they do not grow faster, indefinitely. This is an important result which may seem counterintuitive. Yet it is implied by the simple Solow model.

To see this, consider Figure 3.8, which illustrates the effect of an increase in the savings rate from *s* to *s'*. The savings–investment schedule shifts upwards while the production function schedule remains unchanged. The new steady-state output–labour and capital–labour ratios are both higher at point *B* than

Table 3.3 Average Annual Growth Rates of GDP per Capita since 1960

	1961–1970	1971–1980	1981–1990	1991–2000	2001–2010
France	4.6	3.1	1.8	1.5	0.5
Japan	8.5	3.3	4.1	0.9	0.8
China	2.4	4.4	7.8	9.3	9.8
India	4.1	0.8	3.4	3.6	6.3
Vietnam			2.3	5.9	6.1

Source: World Development Indicators, The World Bank.

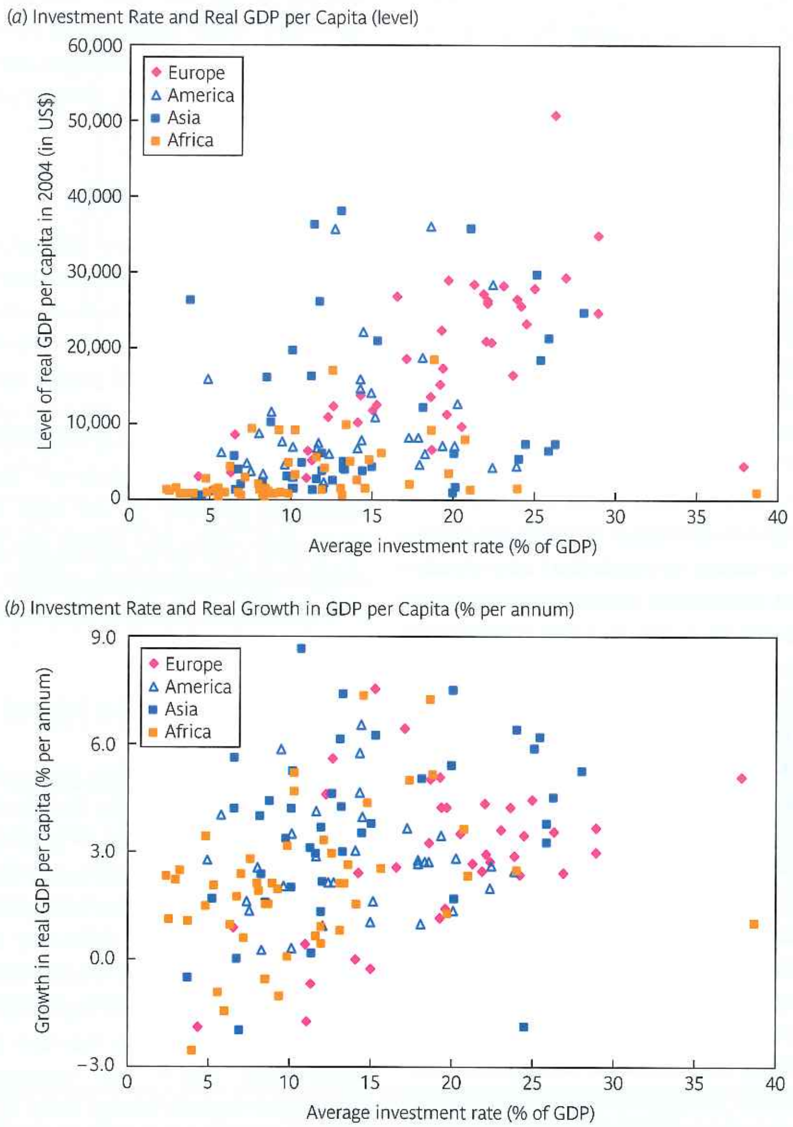


Fig. 3.7 Investment, GDP per Capita, and Real GDP Growth

For a sample of 174 countries over the period of 1950–2004, the correlation coefficient between the investment rate (the ratio of investment to GDP) and the average per capita GDP over the period is high and positive (0.51). The correlation of the investment rate in the countries with real GDP growth is also positive but less striking (0.31).

Source: Penn World Table Version 6.2 September 2006.

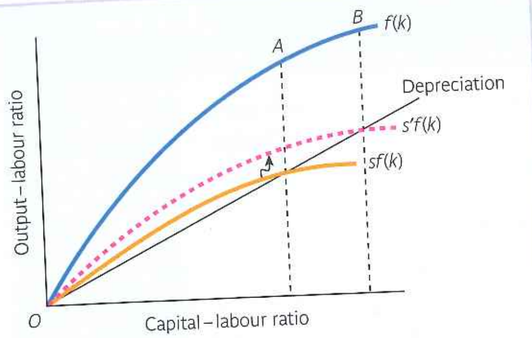


Fig. 3.8 An Increase in the Savings Rate

An increase in the savings rate raises capital intensity (*k*) and the output–labour ratio (*y*).

they were at point *A* beforehand. It will take time, however, for the economy to reach that new steady state. At the initial steady-state position (point *A*), an increase in the saving state causes gross investment to rise. Since depreciation is unchanged, net investment must be positive. The capital–labour ratio starts rising, which raises the output–labour ratio. This will go on until the new steady state is reached at point *B*. During this interim period, therefore, growth is higher, which can give the impression that higher investment rates cause higher economic growth. The boost is only temporary; once the steady state has been reached, no further growth effect can be expected from a higher savings rate. We still need a story to explain growth in output per capita. This is the story told in Sections 3.4 and 3.5.

It may be surprising that increased savings does not affect long-run growth. That result is an implication of diminishing returns. An increase in savings causes the capital stock to rise, but this also means that more capital depreciates and thus needs to be replaced. Increasing amounts of gross investment are needed just to keep the capital stock constant at its higher level. Yet the resources for that increased investment are not sufficiently forthcoming because the marginal productivity of capital decreases. Further additions to the capital–labour ratio yield smaller and smaller increases in income, and therefore in savings and investment, while depreciation rises proportionately with the capital stock. Put simply, decreasing marginal productivity implies that, at some point, saving more is simply not worth it.[14]

3.3.5 The Golden Rule

Figure 3.8 contains an important message: to become richer, you need to save and invest more. But is being richer—in the narrow sense of accumulating capital goods—always necessarily better? Saving requires the sacrifice of giving up some consumption today against the promise of higher income tomorrow, but does saving more today always mean more consumption tomorrow? The answer is not necessarily positive. To see why, note that in the steady state, when the capital stock per capita is $\bar{k}$, savings equal depreciation and the steady-state level of consumption $\bar{c}$ (the part of income that is not saved) is given by:

$$\bar{c} = \bar{y} - s\bar{y} = f(\bar{k}) - \delta\bar{k}. \tag{3.10}$$

In Figure 3.9, consumption per capita is given by the vertical distance between the production function and the depreciation line.[15] If we could choose any saving rate we wanted, we could effectively pick any point of intersection of the savings schedule with the depreciation line, and therefore any level of consumption we so desired. Figure 3.9 shows that consumption is highest at the capital stock for which the slope of the production function is parallel to the depreciation line.[16] The corresponding optimal steady-state capital–labour ratio is indicated as $\bar{k}'$. Now remember that the slope of the production function is the marginal productivity of capital (MPK) while the slope of the depreciation schedule is the rate of depreciation δ. We have just shown that the maximal level of consumption is achieved when the slope of the production function, the marginal product of capital, is equal to the slope of the depreciation line:

$$\text{MPK} = \delta. \tag{3.11}$$

This condition is called the **golden rule**, and can be thought of as a recipe for achieving the highest consumption per capita, given existing technological capabilities. In this case, with no population growth and no technical progress, the golden rule states that the economy maximizes steady-state consumption when the marginal gain from an additional unit of GDP saved and invested in capital (MPK) equals the depreciation rate.

[14] In Chapter 4, we will see that the outcome is very different when the marginal productivity of capital is not decreasing in capital.

[15] Note that everything, including consumption and saving, is measured as a ratio to the labour input, person-hours. As already noted, if the number of hours worked does not change, the ratios move exactly as per capita consumption, savings, output, etc.

[16] An exercise asks you to prove this assertion.

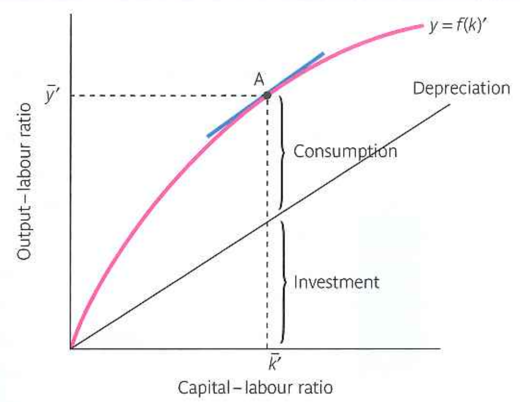


Fig. 3.9 The Golden Rule

Steady-state consumption $\bar{c}$ (as a ratio to labour) is the vertical distance between the production function and the depreciation line. It is at a maximum at point *A* corresponding to the situation where the slope of the production function, the marginal productivity of capital, is equal to δ, the slope of the depreciation line.

What are the consequences of 'disobeying' the golden rule? If the capital–labour ratio exceeds $\bar{k}'$, too much capital has been accumulated, and the MPK is lower than the depreciation rate δ. By reducing savings today, an economy can actually *increase* per capita consumption, both today and in the future. This looks like a free lunch, and indeed, it is one. We say that the economy suffers from **dynamic inefficiency**. Dynamically inefficient economies simply save and invest too much—in order to maintain a capital stock which is simply too large—and consume too little as a result.

A different situation arises if the economy is to the left of $\bar{k}'$. Here, steady-state income and consumption per capita may be raised by saving more, but not immediately; consumption can be increased only in the long run, after the adjustment has occurred. No free lunch is immediately available. More consumption in the future—in the steady state—must be 'earned' by increased saving and reduced consumption at the outset. Moving towards $\bar{k}'$ from a position on the left requires current generations to sacrifice so future generations can enjoy more consumption which will result from more capital and income in the steady state. An economy in such a situation is called **dynamically efficient** because it is not possible to do better without paying the price for it.

The difference between dynamically efficient and inefficient savings rates is illustrated in Figure 3.10,

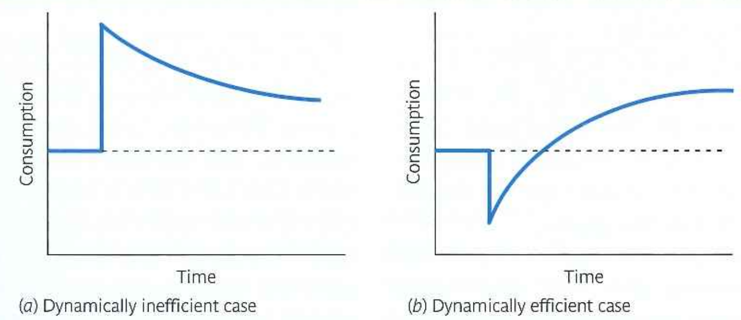


Fig. 3.10 Raising Steady-State Consumption

In a dynamically inefficient economy (*a*), it is possible to permanently raise consumption by reducing saving. In a dynamically efficient economy (*b*), higher future consumption requires early sacrifices.

which shows how we move from one steady state to another one with higher consumption. In the dynamically inefficient case (*a*), it is possible to permanently raise consumption by consuming more now and during the transition to the new steady state. In the dynamically efficient case (*b*), a higher steady-state level consumption is not free and implies a transitory period of sacrifice.

Dynamic inefficiency—in short, too much capital—may have characterized some of the centrally planned economies of Central and Eastern Europe. We say 'may' because the proof that an economy is dynamically inefficient lies in showing that the marginal productivity of its capital lies below the depreciation rate, and neither of these is easily measurable. What we do know is that communist leaders often boasted about their economies' high investment rates, which were in fact considerably higher than in the capitalist West. Yet overall standards of living were considerably lower than in market economies, and consumer goods were in notoriously short supply. Box 3.4 presents the case of Poland.

In dynamically efficient economies, future generations would benefit from raising saving today, but those currently alive would lose. Should governments do something about it? Since it would represent a transfer of revenues from current to future generations, there is no simple answer. It is truly a deep political choice with no solution since future generations don't vote today. A number of factors influence savings, such as taxation, health and retirement systems, cultural norms, and social custom. Importantly, too, saving and investment are influenced by political conditions. Political instability and especially wars, civil or otherwise, can lead to destruction and theft of capital, and hardly encourage thrifty behaviour. As we discuss in Chapter 4, in many of the world's poorest countries, property rights are under constant threat or non-existent.

Box 3.4 Dynamic Inefficiency in Poland?

From the period following the Second World War until the early 1990s, Poland was a centrally planned economy. Savings and investment decisions for the Polish economy were taken by the ruling communist party. The panels of Figure 3.11 compare Poland with Italy, a country with one of the highest saving rates in Europe. The first graph shows the increase in GDP per capita between 1980 and 1990 (the GDP measure is adjusted for purchasing power to take into account different price systems). While Italy's income grew by 25%, Poland's actually shrank by about 5%. The second graph shows the average proportion of GDP dedicated to saving over the same period. Clearly, Poland saved a lot, but received nothing for it in terms of income growth.

The third and fourth panels of Figure 3.11 show that the situation was reversed after 1991, when Poland introduced free markets and abandoned central planning. From 1991 to 2004, per capita GDP increased by 68%, with a lower investment rate than Italy's (which grew by 17%). However, our theory predicts that savings affect the steady-state level of GDP per capita, not its growth rate. In 1980, Poland invested 21.6% of its GDP. By 1990 this rate had fallen to 18.3%. In the period 1990–2004 per capita consumption rose in Poland from \$2,908 to \$7,037, an increase of 142%, compared with 61% in Italy over the same period. Is this proof of dynamic inefficiency, i.e. that a significant part of savings was used merely to keep up an excessively large stock of capital? Anecdotal evidence would suggest so. Stories of wasted resources were common in centrally planned economies: uninstalled equipment rusting in backyards, new machinery prematurely discarded for lack of spare parts, tools ill-adapted to factory needs, etc. One important cause of wastage was a reward system for factory managers. These were often based on spending *plans*, and not on actual output. An alternative interpretation is that the investment was in poor quality equipment, which could not match western technology. No matter how we look at it, savings were not put to their best use in centrally planned Poland.

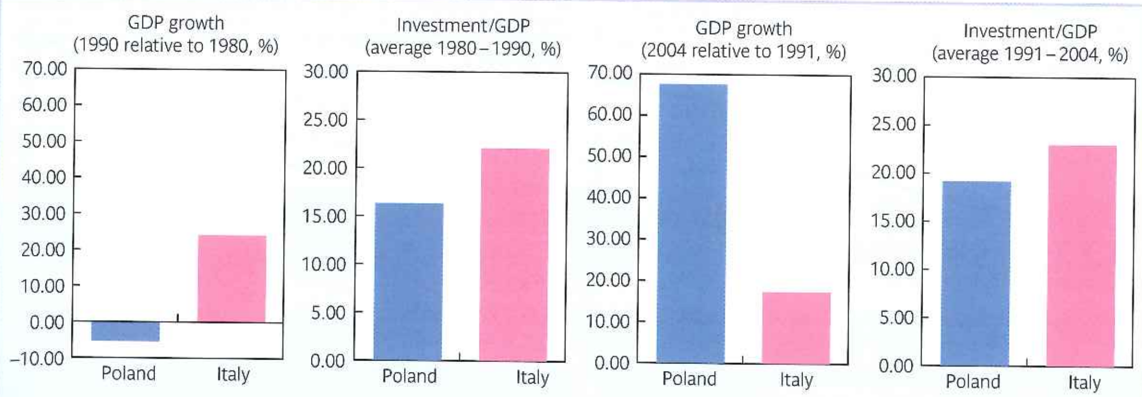


Fig. 3.11 Was Centrally Planned Poland Dynamically Inefficient?

Despite a high investment and savings rate, Polish per capita GDP shrank during the period 1980–1990 while Italy's grew. During the transition period, Poland grew much faster, with a lower investment rate than in Italy.

Source: Heston, Summers, and Aten (2006).

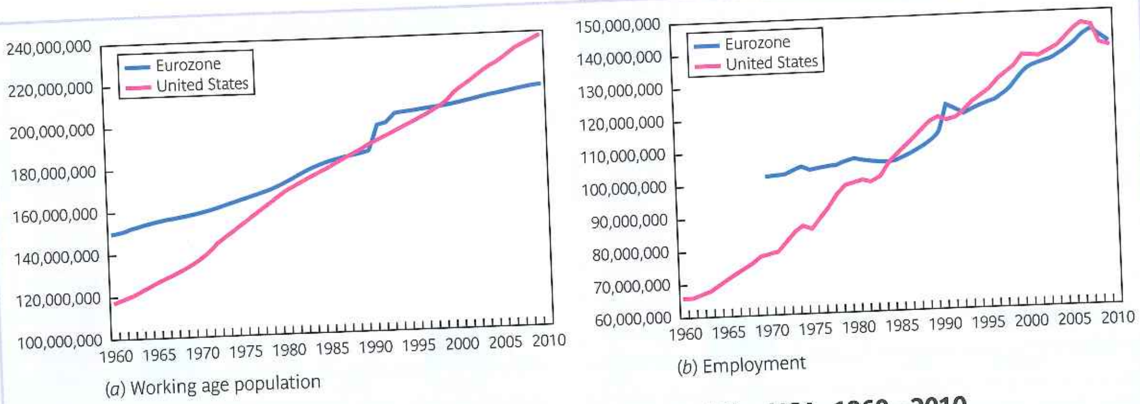
3.4 Population Growth and Economic Growth

A major shortcoming of the previous section is that it does not explain permanent, sustained growth, our very first stylized fact. Capital accumulation, we saw, can explain high living standards and growth during the transition to the steady state but the law of diminishing returns ultimately kicks in. Clearly, some crucial ingredients are missing. One of them is population growth, more precisely, growth in the employed labour force. This section shows that, once we introduce population growth, sustainable long-run growth of both output and the capital stock is possible.

Recall that labour input (person-hours) grows either if the number of people at work increases, or if workers work more hours on average. Later on in this chapter and Chapter 5, we will see that the number of hours worked per person has declined steadily over the past century and a half. Figure 3.12 shows that, despite this fact, employment has been rising, either because of natural demographic forces (the balance between births and deaths), increasing labour force participation (especially by women) or immigration. Overall, more people are at work but they work shorter hours, so the balance of effects is ambiguous. Because the number of hours worked per person cannot and does not rise without bound, we will treat it as constant. Then any change in person-hours is due to exogenous changes in the population and employment, and output per person-hour changes at the same rate as output per capita.

Even though population and employment are growing, the fundamental reasoning of Section 3.3 remains valid: the economy gravitates to a steady state at which the capital–labour and output–labour ratios ($k = K/L$ and $y = Y/L$) stabilize. If L grows at an exogenous rate n, output Y and capital K will also grow at rate n. The relentless increase in the labour input will be the driver of growth in this case. Quite simply, if income per capita is to remain unchanged in the steady state, income must grow at the same rate as the number of people.

The role of saving and capital accumulation remains the same as in the previous section, with



(*a*) Working age population (*b*) Employment

Fig. 3.12 Population and Employment in the Eurozone and the USA, 1960–2010

Population of working age (between 15 and 64) has been growing both in the USA and the countries of the Eurozone, the part of the European Union that uses the euro. Employment has also been growing, albeit less fast in Europe. Note the jump in the euro area in 1991, the year after German unification.

Source: OECD, *Economic Outlook*.

only a small change of detail. The capital accumulation condition (3.9) now becomes:[17]

(3.12) $$\Delta k = sf(k) - (\delta + n)k.$$

The difference is that, for the capital–labour ratio to increase, gross investment must not just compensate for depreciation, it must also provide new workers with the same equipment as those already employed. This process is called **capital-widening** and it explains the presence of n in the last term of the equation.

The situation is presented in Figure 3.13. The only difference with Figure 3.5 is that the depreciation line δk has been replaced by the steeper **capital-widening line** $(\delta + n)k$. The fact that the capital-widening line is steeper than the depreciation line captures the greater need to save when more workers are being equipped with productive capital. The steady state occurs at point A_1, the intersection of the saving schedule and the capital-widening line. At this intersection $\bar{k}_1$, savings are just enough to cover the depreciation and the needs of new workers, so $\Delta k = 0$.

The role of population growth can be seen by studying the effect of an increase in the rate of population growth, from n_1 to n_2. In Figure 3.13 the capital-widening line becomes steeper and the new steady state at point A_2 is characterized by a lower capital–labour ratio $\bar{k}_2$. This result makes sense, since we assume that savings behaviour is unchanged and yet we need more gross investment to equip the constant inflow of new workers. The solution is to provide each worker with less capital. Of course, a lower $\bar{k}$ implies a lower output–labour ratio $f(\bar{k})$. The Solow model with population growth implies that, all other things equal, countries with a rapidly growing population will tend to be poorer than countries with lower population growth. Box 3.5 examines whether it is indeed the case that high population growth lowers GDP per capita.

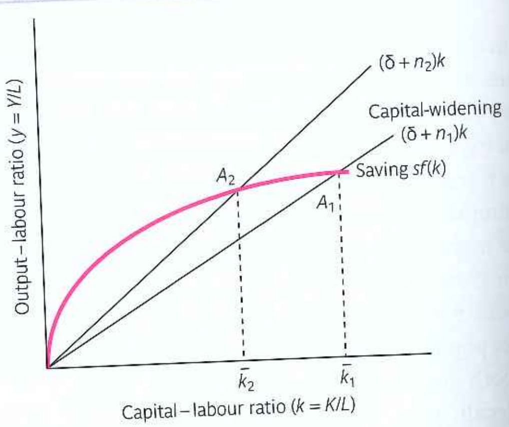


Fig. 3.13 The Steady State with Population Growth

The capital–labour ratio remains unchanged when investment is equal to $(\delta + n_1)k$. This occurs at point A_1, the intersection between the saving schedule $sf(k)$ and the capital-widening line $(\delta + n_1)k$. An increase in the rate of growth of the population from n_1 to n_2 is shown as a counter-clockwise rotation of the capital-widening line. The new steady-state capital–labour ratio declines from $\bar{k}_1$ to $\bar{k}_2$.

[17] The proof requires some calculus based on the principles presented in Box 6.3. The change in capital per capita obeys $\Delta k/k = (\Delta K/K) - (\Delta L/L)$. After substituting $\Delta K = I - \delta K$ and $\Delta L/L = n$ and setting $I = sY$, the equation simplifies to $\Delta k = sy - \delta k - nk = sf(k) - (\delta + n)k$.

Box 3.5 Population Growth and GDP per Capita

Figure 3.14 plots GDP per capita in 2008 and the average rate of population growth over the period 1960–2009. The figure could be seen as confirming the negative relationship predicted by the Solow growth model—and might even be interpreted as support for the hypothesis that population growth impoverishes nations. Thomas Malthus, a famous nineteenth-century English economist and philosopher, claimed exactly this: population growth causes poverty. He argued that a fixed supply of arable land could not feed a constantly increasing population and that population growth would ultimately result in starvation. He ignored technological change, in this case the green revolution which significantly raised agricultural output in the last half of the twentieth century. As we confirm in Section 3.5, technological change can radically alter the outlook for growth and prosperity.

Yet the pseudo-Malthusian view has been taken seriously in a number of less-developed countries, which have attempted to limit demographic growth. The most spectacular example is China, which has pursued a one-child-only policy for decades. At the same time, we need to be careful with simple diagrams depicting relationships between two variables. Not only do other factors besides population growth influence economic growth, but it may well be that population growth is not exogenous. Figure 3.14 could also be read as saying that as people become richer, they have fewer children. There exists a great deal of evidence in favour of this alternative interpretation.

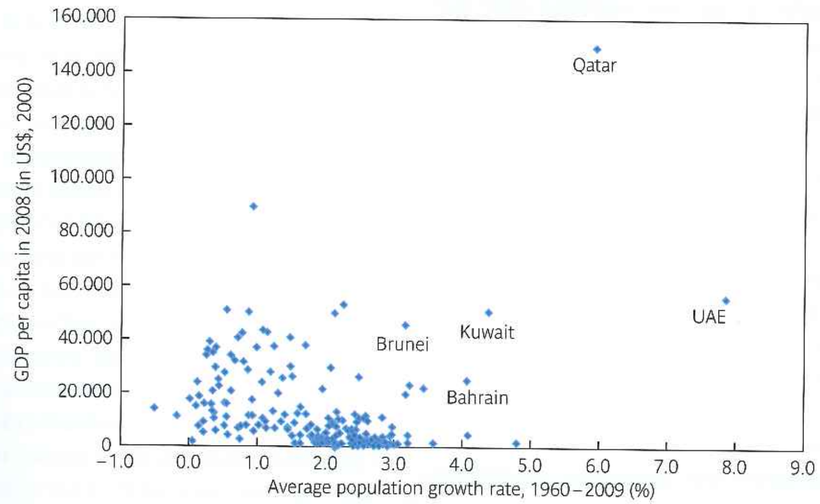


Fig. 3.14 Population Growth and GDP per Capita, 1960–2009

The figure reports data on real GDP per capita and average population growth for 190 countries over almost a half century. The plot indicates a discernible negative association between GDP per capita and population growth, especially when the rich oil-producing countries (United Arab Emirates, Qatar, Kuwait, Bahrain, Oman, Brunei, and Saudi Arabia) are excluded. The sharp population growth observed in these countries is largely to due to immigration.

Source: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

At what level of investment does an economy with population growth maximize consumption per capita? Because the number of people who are able to consume and save is growing continuously, the golden rule must be modified accordingly. Following the same reasoning as in Section 3.3, we note that steady-state investment per person-hour is $(\delta + n)\bar{k}$, so consumption per person-hour $\bar{c}$ is given by $f(\bar{k}) - (\delta + n)\bar{k}$. Proceeding as before, it is easy to see that consumption is at a maximum when:

$$\text{MPK} = \delta + n. \tag{3.13}$$

The 'modified' golden rule equates the marginal productivity of capital with the sum of the depreciation rate δ and the population growth rate n. The intuition developed here continues to apply: the marginal product of an additional unit of capital (per capita) is set to its marginal cost, which now includes not only depreciation, but also the capital-widening investment necessary to equip future generations with the same capital per head as the current generation. A growing population will necessitate a higher marginal product of capital at the steady state. The principle of diminishing marginal productivity implies that the capital–labour ratio must be lower. Consequently, output per head will also be lower.

3.5 Technological Progress and Economic Growth

Taking population growth into account gives one good reason why output and the capital stock can grow permanently, and at the same rate. While this satisfies Kaldor's second stylized fact, the picture remains incomplete: so far, we conclude that the capital–labour and output–labour ratios are *constant* in the steady state. That standards of living are not rising is inconsistent with Kaldor's first stylized fact and the data reported in Table 3.1. We still need to account for the fact that per capita income and capital stock grow, *and* grow at the same rate.

As it turns out, we have ignored **technological or technical progress**. Over time, increased knowledge and better, more sophisticated techniques make workers and the equipment they work with more productive. With a slight alteration, our framework readily shows how technological progress can explain ever-rising living standards. To do so, once more, we reformulate the **aggregate production function** introduced in (3.1). Technological progress means that more output can be produced with the same quantity of equipment and labour. The most convenient way to think of technology is as an additional factor of production. In the production function, we add a measure of the state of technology, A, which raises output at given levels of capital stock and employment:

$$Y = F(\underset{+}{A}, \underset{+}{K}, \underset{+}{L}). \tag{3.14}$$

When A increases, Y rises, even if K and L remain unchanged. For this reason, A is frequently called **total factor productivity**. Yet A is *not* really a factor of production, but rather a method or technique of production. No firm pays for knowledge; each firm just benefits from it. It is best thought of as 'best practice', which is assumed to be available freely to all. Innovations of all sorts continuously raise productivity. At this point, it will be convenient to assume that A increases at a constant rate a, without trying to explain how and why. Technological progress, which is the increase in A, is therefore considered as exogenous.

It turns out that it is possible to relate our analysis to previous results in this chapter in a straightforward way. First, we modify (3.14) to incorporate technical progress in the following particular way:

$$Y = F(K, AL). \tag{3.15}$$

In this formulation, technological progress acts directly on the effectiveness of labour. (For this reason it is sometimes called *labour-augmenting* technical progress.)[18] An increase in A of, say, 10% has the same impact as a 10% increase in employment, even though the number of hours worked hasn't changed. The term AL is known as **effective labour** (or labour in efficiency units) to capture the idea that, with the same equipment, one hour of work today produces more output than before because A is higher. Effective labour AL grows for two reasons: (1) more labour L, and (2) greater effectiveness A. For this reason, the rate of growth of AL is now given by $a + n$.

Now we change the notation a little bit. We redefine y and k as ratios of output and capital *relative to effective labour*: $y = Y/AL$, $k = K/AL$. Once this is done, it is only a simple step to recover the now-familiar production function in intensive form, $y = f(k)$.[19] Not surprisingly, the ratio of capital to effective labour evolves almost exactly as before:

$$\Delta k = sf(k) - (\delta + a + n)k. \tag{3.16}$$

The reasoning is the same as when we introduced population growth. There we noted that, to keep the capital–labour ratio K/L constant, the capital stock K must rise to compensate for both depreciation (δ) and population growth (n). Now we find that, to keep the capital–effective labour ratio $k = K/AL$ constant, the capital stock K must also rise to keep up with workers' enhanced effectiveness (a). So k will increase if saving $sf(k)$, and hence gross investment, exceeds the capital accumulation needed to make up for depreciation δ, population growth n, and increased effectiveness a. From there on, it is a simple matter to modify Figure 3.13 to draw Figure 3.15. The steady state is now characterized by constant ratios of capital and output to effective labour ($y = Y/AL$ and $k = K/AL$).

Constancy of these ratios in the steady state is a very important result. Indeed, if Y/AL is constant, it means that Y/L grows at the same rate as A. If the average number of hours remains unchanged, then income per capita must grow at the rate of technological progress, a. In other words, we have finally uncovered the explanation of Kaldor's first stylized fact: the continuous increase in standards of living is due to technological progress. Since K/AL is also constant, the capital stock per capita also grows in the steady state at the same rate as A, which is Kaldor's second stylized fact.

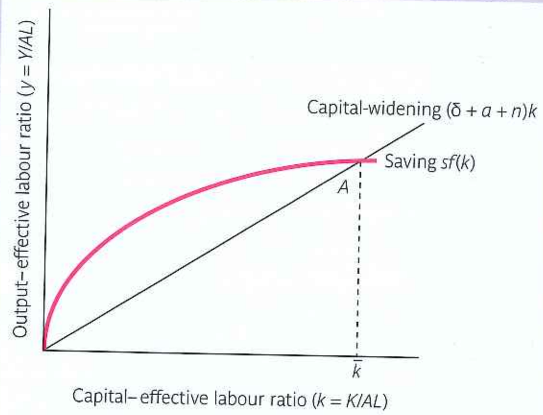


Fig. 3.15 The Steady State with Population Growth and Technological Progress

In an economy with both population growth and technological progress, inputs and output are measured in units per effective labour input. The intensive form production function inherits this property. The slope of the capital accumulation line is now $\delta + a + n$, where a is the rate of technological progress. The steady state occurs when investment is equal to $(\delta + a + n)k$ (point A), which is the intersection of the saving schedule $sf(k)$ with the capital-widening line $(\delta + a + n)k$. At the steady-state A, output and capital increase at the rate $a + n$, while GDP per capita increases at the rate a.

Figure 3.16 illustrates these results. Because of diminishing marginal productivity, capital accumulation alone cannot sustain growth. Population growth explains GDP growth, but not the sustained increase of standards of living over the centuries. Technological progress is essential for explaining economic growth in the long run. Rather than creating misery in the world, it turns out to be central to improvements in standards of living.

Note that an increase in the rate of technological progress, a, makes the capital-widening line steeper

[18] Obviously, technological innovation sometimes also makes capital more productive. It is much simpler to assume that technological progress is labour-augmenting and it is enough to provide a reasonable—if incomplete—interpretation of the growth phenomenon.

[19] Constant returns to scale implies that $y = F(K, AL)/AL = F(K/AL, 1)$.

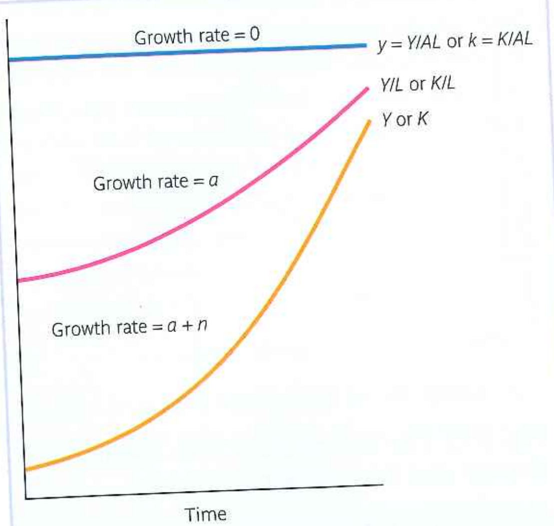


Fig. 3.16 Growth Rates along the Steady State

While output and capital measured in effective labour units (Y/AL and K/AL) are constant in the steady state, output–labour and capital–labour ratios (Y/L and K/L) grow at the rate of technological progress a, and output and the capital stock (Y and K) grow at the rate $a + n$, the sum of the rates of population growth and technological progress.

than before. In Figure 3.15 this would imply lower steady-state ratios of capital and output to effective labour. This does not mean that more rapid technological progress is a bad thing, because what counts is income per person (Y/L), not income per unit of effective labour (Y/AL). The latter is just an accounting device to help describe a dynamic economy analytically and graphically. In the steady state, the higher a is, the faster Y/L and standards of living will grow.

The discussion can be extended in a natural way to address the issue of the golden rule. Redefining c as the ratio of aggregate consumption (C) to effective labour (AL), the following modified version of (3.10) will hold in the steady state:

$$\bar{c} = f(\bar{k}) - (\delta + a + n)\bar{k}.$$

The modified golden rule now requires that the marginal productivity of capital be the sum of the rates of depreciation, of population growth, and of technological change:

$$\text{MPK} = \delta + a + n. \tag{3.17}$$

Maximizing consumption per capita is equivalent to making consumption per unit of effective labour as large as possible. To do this, an economy now needs to invest capital per effective unit of labour to the point at which its marginal product 'covers' the investment requirements given by technical progress (a), population growth (n), and capital depreciation (δ).

3.6 Growth Accounting

3.6.1 Solow's Decomposition

As shown in Figure 3.16, we have now identified three sources of GDP growth: (1) capital accumulation, (2) population growth, and (3) technological progress. It is natural to ask how large the contributions of these factors are to the total growth of a nation or a region. Unfortunately, it is difficult to measure technological progress. Computers, for instance, probably raise standards of living and growth, but by how much? Some people believe that the 'new economy', brought on by the information technology revolution, will push standards of living faster than ever. Others are less optimistic that the effect is any larger than other great discoveries which mark economic history. Box 3.6 provides some details on this exciting debate.

Robert Solow, who developed the theory presented in the previous sections, devised an ingenious method of quantifying the extent to which technological progress accounts for growth. His idea was to start with the things we can measure: GDP growth, capital accumulation, and hours worked. Going back to the general form of the production function (3.14), we can measure output Y and two inputs, capital K and labour L. Once we know how much GDP has increased, and how much of this increase is explained by capital and hours worked, we can interpret what is left, called the **Solow residual**, as due to the increase in A, i.e. $a = \Delta A/A$:

$$\text{Solow residual} = \frac{\Delta Y}{Y} - \text{output growth due to growth in capital and hours worked.}$$[21]

We now track down the **Solow decomposition**.

3.6.2 Capital Accumulation

Table 3.4 shows that, typically, capital has been growing at about 3–5% per year over most of the twentieth century in the developed countries. Capital accumulation accelerated sharply in the 1950s and 1960s as part of the post-war reconstruction. Many European countries accumulated capital considerably faster than the USA and the UK up until the mid-1970s, the reason being that continental Europe was poorer at

Box 3.6 The New Economy: Another Industrial Revolution?

The striking changes brought about by the ICT (information and communications technologies) revolution, which include the internet, wireless telecommunications, MP3 players, and the conspicuous use of electronic equipment, have led many observers to conclude that a new industrial revolution is upon us. Figure 3.17 reports estimates of overall increases in multifactor productivity in the USA, computed as annual averages over four periods. A difference of 1% per year cumulates to 28% after 25 years. The figure shows a formidable acceleration in the period 1913–1972, and again over 1995–1999; hence the case for a second industrial revolution.

While initially there was much scepticism about the true impact of the ICT revolution—Robert Solow himself said early on that 'computers can be found everywhere except in the productivity statistics'—there is compelling evidence that ICT have indeed deeply impacted the way we work and produce goods and services, and have ultimately increased our standards of living significantly. These total factor productivity gains, according to recent research by Kevin Stiroh, Dale Jorgenson, and others, can be found both in ICT-*producing* as well as ICT-using sectors.[20] Strong gains in total factor productivity have been observed in the organization of retail trade, as well as in manufacturing and business services. Even more interesting is the fact that not all economies around the world have benefited equally from productivity improvements measured in the USA. In particular, some EU countries continue to lag behind in ICT adoption as well as innovation.

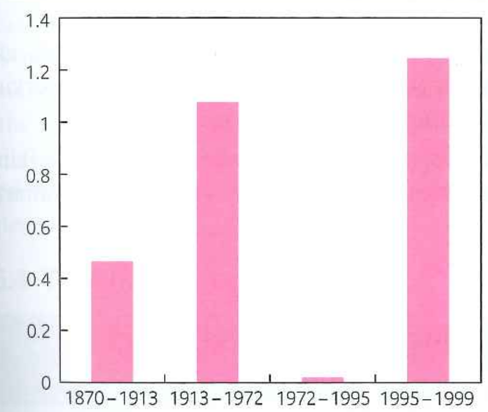


Fig. 3.17 Multifactor Productivity in the USA (average annual growth, %)

The average annual increase in A (multifactor productivity) accelerated sharply after 1913, almost came to a halt in the years 1972–1995, and seems to have recovered vigorously at the end of the 1990s.

Source: Gordon (2000).

20 See the references at the end of the book.

21 Formally, the Solow residual is $\frac{\Delta A}{A} = \frac{\Delta Y}{Y} - \left[(1-s_L)\frac{\Delta K}{K} + s_L\frac{\Delta L}{L}\right]$, where s_L is the labour share, defined as the share of national income paid to labour in the form of wages and non-wage compensation, and $1 - s_L$ is the income share of capital. The WebAppendix shows how this formula can be derived from the production function.

Table 3.4 Growth of Real Gross Fixed Capital Stock, 1913–2010 (% per annum)

	1913–50	1950–73	1973–87	1987–2010
France	1.2	6.4	3.7	3.3
Germany	1.1	7.7	2.7	1.9*
Netherlands	2.4	6.9	2.2	2.5
UK	1.6	5.7	2.3	3.6
USA	1.7	3.8	2.6	2.8

* 1991–2008.

Sources: Maddison (1991); OECD, *Economic Outlook*.

the end of the Second World War. These sustained periods of rapid capital accumulation fit well the description of catch-up, when the capital stock is below its steady-state level.

3.6.3 Employment Growth

The most appropriate measure of labour input is total number of hours worked. For several reasons, however, growth in population or the number of employees does not necessarily translate into increased person-hours. To understand this, we can rewrite the total number of hours in the following way:

$$\text{Total hours worked} = (\text{hours/employee}) \times (\text{employee/population}) \times \text{population}.$$

The total number of hours worked can increase for three reasons:

- Obviously, population growth. Everything else unchanged, more people provide more working hours. But many things can and do change.
- The proportion of people who work. Some working-age people are unemployed and others voluntarily stay out of work for various reasons. In addition, people live longer, study longer, and retire earlier. Furthermore, women have increased their labour force participation over the past 30 years.[22]
- Hours worked per person. Over time, those who work tend to work fewer hours per day and fewer days per year.

Table 3.5 Population, Employment, and Hours Worked, 1913–2010

	Population growth (% per annum)	Employment growth (% per annum)	Growth in hours worked per person (% per annum)	Hours worked per person in 1913	Hours worked per person in 2010
France	0.4	0.3	−0.5	2,588	1,561
Germany	0.5	0.7	−0.5	2,584	1,418
Netherlands	1.1	1.3	−0.5	2,605	1,372
UK	0.5	0.5	−0.4	2,624	1,647
USA	1.5	1.6	−0.4	2,605	1,690
Japan	0.9	0.9	−0.4	2,588	1,713

Sources: Maddison (2006); Groningen Growth and Development Centre and the Conference Board, Total Economy Database, January 2011.

Table 3.5 shows that these effects have roughly cancelled each other out so that, in the end, employment and population size have increased by similar amounts in our sample of developed countries.

Table 3.5 also documents the sharp secular decline in the number of hours worked per person in the developed world. The long-run trend is a consequence of shorter days, shorter workweeks, fewer weeks per year, and fewer years worked per person. This is why the number of person-hours has declined across the industrial world. The dramatic decline in hours worked per person is a central feature of the growth process; an average annual reduction of 0.5% per year means a total decline of 40% over a century. As societies become richer, demand for leisure increases. The last two columns of the table reveal a massive jump in leisure time available, which is as important a source of improvement of human welfare as increases in material wealth.

3.6.4 Technological Change

Table 3.6 presents the Solow decomposition in two different contexts. The first employs historical data for the period 1913–1987. The second examines the same countries over the last two decades. We see that growth in inputs of labour and capital account for only one-half to two-thirds of total economic growth in these economies. The rest is the Solow residual, and confirms the importance of technological progress. Technological change appears to be remarkably steady, contributing to growth about 1% per year. We see a tendency to an acceleration of technological change since the late 1980s, with the notable exception of the UK, which makes a comeback over the last decade.

Table 3.6 The Solow Decomposition (average annual growth rates)

(*a*) 1913–1987*

Country	GDP	Contribution of inputs	Residual
France	2.6	1.1	1.0
Germany	2.8	1.4	0.8
Netherlands	3.0	2.0	0.4
UK	1.9	1.2	0.5
USA	3.0	2.0	0.7
Japan	4.7	3.0	0.5

* An adjustment is made to account for the modernization of productive capital.

Source: Maddison (1991:158).

(*b*) 1987–1997

	GDP	Contribution of inputs	Residual
France	2.0	1.1	1.0
Germany*	1.4	0.2	1.2
Netherlands	2.9	1.8	1.1
UK	2.2	1.4	0.7
United States	3.0	2.5	0.5
Japan	2.7	1.4	1.3

* 1991–1997.

(*c*) 1997–2006

	GDP	Contribution of inputs	Residual
France	2.2	1.3	1.0
Germany	1.4	0.6	0.8
Netherlands	2.3	1.4	0.9
UK	2.7	1.7	1.0
United States	3.0	1.9	1.1
Japan	1.2	0.1	1.1

Sources: Maddison (1991); authors' calculation based on Groningen Growth and Development Centre and the Conference Board, Total Economy Database, January 2007, available at <www.ggdc.net>; OECD, *Economic Outlook*.

[22] Chapter 5 explores these and related issues in more detail.