

CHAPTER 3

Economic Growth

3.1 Introduction

Of all the issues facing development economists, economic growth has to be one of the most compelling. In Chapter 2, we noted the variety of growth experiences across countries. It is true that all of these numbers, with very few exceptions, were in the single digits, but we also appreciated the power of exponential growth. A percentage point added or subtracted can make the difference between stagnation and prosperity over the period of a generation. Small wonder, then, that the search for key variables in the growth process can be tempting. For precisely this reason, the theory and empirics of economic growth (along with the distribution of that growth) has fired the ambitions and hopes of academic scholars and policy makers alike. I was certainly inspired by Robert Lucas's Marshall Lectures at the University of Cambridge:

Rates of growth of real per-capita income are diverse, even over sustained periods. Indian incomes will double every 50 years; Korean every 10. An Indian will, on average, be twice as well off as his grandfather; a Korean 32 times.¹

I do not see how one can look at figures like these without seeing them as representing *possibilities*. . . 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.

¹As we have seen, this is no longer true of India and Korea post 1990, but the general point, made at a time when India was growing slowly, is still valid.

Never mind the fact that India has grown at far faster rates since these words were penned. The sentiment still makes sense: Lucas captures, more keenly than any other writer, the passion that drives the study of economic growth. We can sense the big payoff, the possibility of change with extraordinarily beneficial consequences, if one only knew the exact combination of circumstances that drives economic growth.

If only one knew... , but to expect a single theory about an incredibly complicated economic universe to deliver that would be unwise. Yet theories of economic growth do take us some way in understanding the development process, at least at an aggregate level. This is especially so if we supplement the theories with what we know empirically.

3.2 Modern Economic Growth: Basic Features

Economic growth, as the title of Simon Kuznets' pioneering book [1966] on the subject suggests, is a relatively "modern" phenomenon. Today, we greet 3% annual rates of per capita growth with approval but no great surprise. But throughout most of human history, such growth — or indeed any growth at all — was the exception rather than the rule. In fact, it isn't far from the truth to say that modern economic growth was born after the Industrial Revolution in Britain.

Consider the growth rates of the world's leaders over the past four centuries. During the period 1580–1820 the Netherlands was the leading industrial nation; it experienced an average annual growth in real GDP per worker hour² of roughly 0.2%. The United Kingdom, leader during the approximate period 1820–90, experienced an annual growth of 1.2%. That's (much) faster than the Netherlands, true, but nevertheless a veritable tortoise compared to today's hares. And since then it's been the United States, but from 1890 to the present it has averaged around 2% a year, dropping to more sedate 1.7% over 1990–2011. That is certainly impressive, but it still isn't what we've seen lately: first from Japan and then from East Asia more generally, and more recently South Asia.

A little calculation suggests that you don't even have to look at history to establish the modernity of economic growth. Simply run our trusty formula on doubling times backwards; (Chapter 2, footnote 15). Let's use what by today's high standards is a pretty moderate number: 3% per year. A country growing at that rate will halve its income every 23 years or so,

²Notice that we are referring here not to growth in per capita GDP as such, but growth in GDP *per worker hour*, or labor productivity. However, the data suggests that the former is largely driven by the latter.

Country	1850	1930	2010	1930/1850	2010/1930
Austria	1,650	3,586	24,096	2.2	6.7
Belgium	1,847	4,979	23,557	2.7	4.7
Canada	1,330	4,811	24,941	3.6	5.2
Denmark	1,767	5,341	23,513	3.0	4.4
Finland	911	2,666	23,290	2.9	8.7
France	1,597	4,532	21,477	2.8	4.7
Germany	1,428	3,973	20,661	2.8	5.2
Japan	681	1,850	21,935	2.7	11.9
Netherlands	2,355	5,603	24,303	2.4	4.3
Norway	956	3,627	27,987	3.8	7.7
Sweden	1,076	4,238	25,306	3.9	6.0
Switzerland	2,293	8,492	25,033	3.7	2.9
United Kingdom	2,330	5,441	23,777	2.3	4.4
United States	1,849	6,213	30,491	3.4	4.9
Simple Average	1,576	4,668	24,312	3.0	5.2

Table 3.1. Per capita GDP (1990 international dollars) in selected OECD countries, 1850–2010.

Source: Maddison [2008] and Bolt and Van Zanden [2013]. First-column Switzerland data from 1851.

which means that running back 200 years, that country would have to have an income around 1/350th of what it has today! For the United States, that would mean an princely annual income of around \$100 per year in 1800. That was most assuredly not the case. And poorer countries extrapolated backwards at this rate would vanish.

But of course, this sort of calculation isn't merely theoretical. You can see the acceleration, *even* among now-developed countries, and even if we go right to 2010 (remembering that the first decade of the 21st century hasn't exactly been a bed of roses). Table 3.1 provides a historical glimpse of the period 1850–2010, and shows how growth has transformed the world. This table displays per capita real GDP (valued in 1990 international dollars) for selected OECD countries in the equally spaced years 1850, 1930, and 2010. The penultimate column gives us the ratio of per capita GDP in 1930, at the peak just before the Depression, to its counterpart in 1850. The last column does the same for the years 2010 and 1930. The numbers are pretty stunning. On average (see the last row of the table), GDP per capita in 1930 was 3 times the figure for 1850, but the corresponding ratio for the equally long period between 1930 and 2010 is by 1978 is 5.2! A nearly

Country	1982	1996	2009	Country	1982	1996	2009
Argentina	34.8	28.8	31.6	India	3.7	4.4	7.2
Bangladesh	2.4	2.2	3.1	Indonesia	5.7	8.6	9.1
Botswana	15.8	22.4	29.1	Malaysia	20.9	28.3	30.5
Brazil	27.6	22.3	22.5	Mexico	42.7	29.7	29.8
Chile	19.4	27.8	31.3	Nigeria	5.4	4.1	4.8
China	2.3	5.8	14.8	Pakistan	5.1	5.5	5.7
Cote d'Ivoire	9.9	5.2	3.7	Rwanda	3.3	1.9	2.5
Egypt	10.5	10.1	12.3	South Africa	34.8	22.0	22.3
Ethiopia	2.3	1.5	2.0	Sri Lanka	6.5	7.3	10.4
Ghana	3.3	2.8	3.4	Thailand	9.5	17.1	17.4

Table 3.2. Per capita GDP in Selected Developing Countries Relative to that of the United States (%), 1982–1996–2009.

Source: *World Bank Development Report* [2011].

sixteen-fold increase in real per capita GDP in the space of 150 years cannot but transform societies completely. The developing world, currently going through its own transformation, will be no exception.

Indeed, in the broader sweep of historical time, the development story has only just begun. In the nineteenth and twentieth centuries, only a handful of countries, mostly in Western Europe and North America, and largely represented by the list in Table 3.1, could manage the “takeoff into sustained growth,” to use a well-known term coined by the economic historian W. W. Rostow. Throughout most of what is commonly known as the Third World, the growth experience only began well into this century; for many of them, probably not until the post-World War II era, when colonialism ended. Although detailed and reliable national income statistics for most of these countries were not available until only a few decades ago, the economically backward nature of these countries is amply revealed in less quantitative historical accounts, and also by the fact that they are way behind the industrialized nations of the world today in per capita GDP levels. To see this, refer to Table 3.2, which records the per capita incomes of several developing countries (and some now-developed countries as well) *relative* to that of the United States, for the last two decades of the twentieth century. I don’t plan to be around to update this table 50 years from now, but I would be very curious to know what it will look like.

The table makes it obvious that despite the very high growth rates experienced by several developing countries, there is plenty of catching-up to do. Moreover, there is a twist in the story that wasn’t present a century

ago. Then, the now-developed countries grew (although certainly not in perfect unison) in an environment uninhabited by nations of far greater economic strength. Today, the story is completely different. The developing nations not only need to grow, they *must* grow at rates that far exceed historical experience. The developed world already exists, and their access to economic resources is not only far higher than that of the developing countries, but *the power afforded by this access is on display*. The urgency of the situation is further heightened by the extraordinary flow of information in the world today. People are increasingly and more quickly aware of new products elsewhere and of changes and disparities in standards of living the world over. Exponential growth at rates of 2% may well have significant long-run effects, but they cannot match the parallel growth of human aspirations, and the increased perception of global inequalities. Perhaps no one country, or group of countries, can be blamed for the emergence of these inequalities, but they do exist, and the need for sustained growth is all the more urgent as a result.

3.3 The Beginnings of a Theory

3.3.1 Savings and Investment. In its simplest terms, economic growth is the result of abstention from current consumption. An economy produces goods and services. The act of production generates income, which in turn is used to buy these goods and services. Exactly *which* goods are produced depends on individual preferences and the distribution of income, but as a broad first pass, the following statement is true: commodity production creates income, which creates the demand for those very same commodities.

Let's go a step further and broadly classify commodities into two groups. We may think of the first group as *consumption goods*, which are produced for the express purpose of satisfying human wants and preferences. The mangos you buy at the market, or a tshirt, or an iPad all come under this category. The second group of commodities consists of what we might call *capital goods*, which we may think of as commodities that are produced for the purpose of producing *other* commodities. A blast furnace, a conveyor belt, or a screwdriver might come under the second category.³

³It should be clear from our examples that there is an intrinsic ambiguity regarding this classification. Although a mango or a blast furnace may be easily classified into its respective category, the same is not true of screwdrivers or an iPad. The correct distinction is between goods that generate current consumption and those that produce future consumption, and many goods embody a little of each.

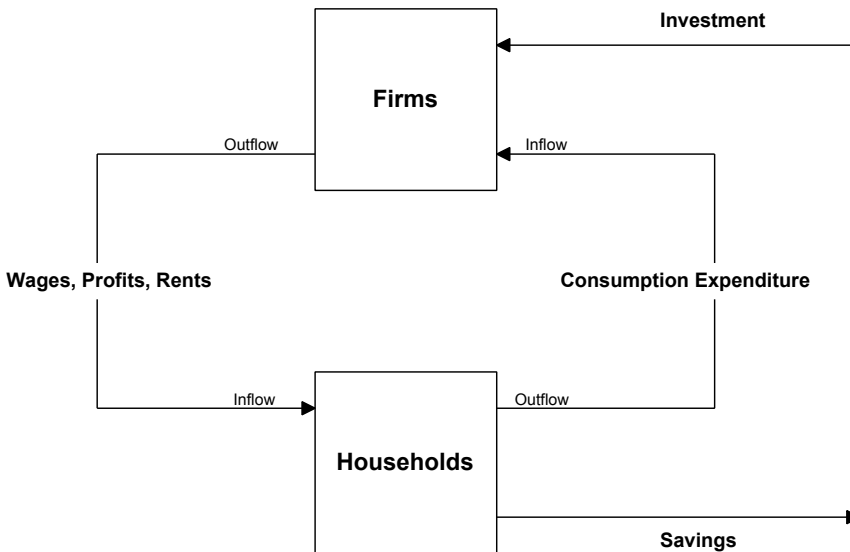


Figure 3.1. Production, consumption, savings, and investment.

Looking around us, it is obvious that the income generated from the production of all goods is spent on *both* consumer goods and capital goods. Typically, households buy consumer goods, whereas firms buy capital goods to expand their production or to replace worn-out machinery. That immediately raises a question: if income is paid out to households, and if households spend their income on consumption goods, where does the money for capital goods come from? How does it all add up?

The answer to this question is simple, although in many senses that we ignore here, deceptively so: households save. No doubt some borrow too, to finance current consumption, but on the whole, national savings is generally positive. All income is not spent on current consumption. By abstaining from consumption, households make available a pool of funds (via deposits, stock purchases, or undeclared dividends) that firms use to buy capital goods. This latter purchase is the act of *investment*. Buying power is channeled from savers to investors through banks, individual loans, governments, and stock markets. How these transfers are actually carried out is a story in itself. Later chapters will tell some of this story.

Implicit in the story is the idea of *macroeconomic balance*. Figure 3.1 depicts a circuit diagram with income flowing “out” of firms as they produce and income flowing back “into” firms as they sell. You can visualize savings as a leakage from the system: the demand for consumption goods alone falls short of the income that created this demand. Investors fill this

gap by stepping in with their demand for capital goods. Macroeconomic balance is achieved when this investment demand is at a level that exactly counterbalances the savings leakage.

By entering a new business, by expanding a current business, or by replacing worn-out capital, investment creates a market demand for *capital goods*. These goods add to the stock of capital in the economy and endow it, in the future, with an even larger capacity for production, and so an economy grows. Note, however, that without the initial availability of savings, it would not be possible to invest and there would be no expansion. This is the simple starting point of all of the theory of economic growth.

3.3.2 Inputs, Outputs and the Production Function. At the heart of it all is, of course, production, that most central of all activities that converts inputs to outputs. It will help — as an organizing device — to introduce the *production function*. It is a simple mathematical description of how various inputs (such as capital, land, labor, and various raw materials) are combined to produce outputs. An easy example is one in which just two inputs — capital and labor — combine to produce a single output. Symbolically, we write

$$Y = F(K, L)$$

where K stands for capital, L for labor, Y for output, and $F(K, L)$ is mathematical notation for some function that converts input pairs (K, L) to output Y . A classic instance is the *Cobb-Douglas* production function, in which we write

$$(3.1) \quad Y = AK^\alpha L^\beta,$$

for some positive constants A , α and β . The parameter A is a measure of the degree of technological proficiency in the economy. The larger it is, the higher is output for any fixed combination of K and L . The parameters α and β capture both the relative importance of each input, as well as whether (and how much) the marginal returns to each input diminish. It is easy to verify that if α lies between 0 and 1, then there are *diminishing returns to capital*: each additional input of capital increases output, but by a progressively smaller amount. (The same is true of labor, if β lies between 0 and 1.) Indeed, “diminishing returns to each input” is a reasonable presumption: if more and more of a particular input is added, *without changing the amounts of any other input*, its marginal efficacy in producing fresh output should go down.

In contrast, if *all* inputs — or overall “scale” for short — are increased in the same proportion, it seems reasonable that output should climb in the same proportion as well. The argument that makes this sound reasonable is one

of pure replication: if you've doubled *every* input, then it's like installing a perfect twin of a manufacturing unit: how can output not double? That typically starts a discussion with some interesting philosophical twists, but we will bypass such matters for the moment.⁴ This phenomenon in which output changes in the same proportion as all inputs is called *constant returns to scale*. It is very easy to verify that in the Cobb-Douglas case described by equation (3.1), "constant returns to scale" is captured by the additional restriction that $\alpha + \beta = 1$.

That isn't to say that the other cases of diminishing returns to scale ($\alpha + \beta < 1$) or increasing returns to scale ($\alpha + \beta > 1$) should be summarily dismissed. We will return to such matters later in the text. But the assumption of constant returns to scale, coupled with diminishing returns to each input, is a good starting point.

When there are constant returns to scale, we can typically express all productive activity in per capita terms; that is, by dividing by the amount of labor being used in production. The Appendix to this chapter shows you how to do this quite generally for production functions, but the Cobb-Douglas case is particularly easy. Change β to $1 - \alpha$ in (3.1) to get

$$(3.2) \quad Y = AK^\alpha L^{1-\alpha},$$

which gives us a familiar functional form much used in the literature, and now divide by L to see that

$$(3.3) \quad y = Ak^\alpha$$

where the lower case letters y and k stand for per capita magnitudes: Y/L and K/L respectively. As before, α , typically a number between 0 and 1, is an inverse measure of diminishing returns to capital. The lower the value of α , the greater the "curvature" of the production function and the higher the extent of diminishing returns.

Figure 3.2 depicts a production function in per capita format. Typically, it will display diminishing returns to per-capita capital. As k increases, so does y (of course), but in a progressively more muted way. Thus the output-capital *ratio* y/k will fall as k climbs, because of a relative shortage of labor. Just how quickly it falls will depend on the extent of diminishing returns to capital.

⁴One question has to do with whether we can really double *all* inputs within the mathematical frame of the production function. What about, say, that elusive input called entrepreneurship or oversight: are we doubling that too? If not, output might less than double. On the other hand, if bricks-and-mortar build a production facility and we double the amount of bricks, the volume of the facility will typically more than double, the former being proportional to surface area and the latter rising at $(\text{surface area})^{3/2}$.

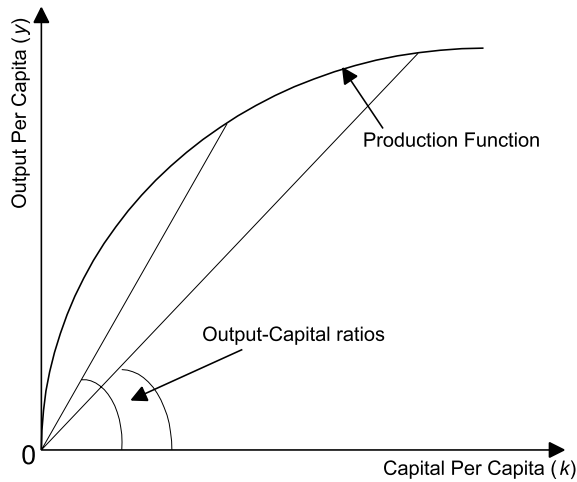


Figure 3.2. The production function in per-capita form.

The concept of a production function already allows to sift and sort some preliminary thoughts about underdevelopment. For instance, using equation (3.1), one might argue that the A term is somehow lower for developing countries: it is in technical knowledge that development lies. Perhaps patent restrictions and other limitations on the flow of knowledge across countries are fundamentally responsible for underdevelopment. Alas, such arguments beg the question of how now-developed countries came to acquire the knowledge they did, and indeed, they ignore the willingness and ability of human beings to take apart new products and simply learn how to rebuild them.

One important interpretation of the “differences in technical knowhow” viewpoint is that developing countries lack a skilled and educated labor force. But properly viewed, these differences should not be thrown into the A -term. Rather, we should recognize that the production functions described in equations such as (3.1) are two-input caricatures of reality. For instance, we could write

$$Y = F(K, \text{labor of different qualities})$$

and recognize that developing countries have lower endowments of some types of labor. That’s conceptually different from saying that they are incapable of producing the same output even with all relevant labor and nonlabor inputs at hand. One quick example using the Cobb-Douglas formulation is to write

$$(3.4) \quad Y = AK^\alpha(eL)^\beta$$

where e is years of schooling per person. That mathematically links up with “lower A ,” but it’s conceptually very different, pointing the finger as it does to education and not technology.

Another interpretation of “different A ” is that while the same technologies are available in both developed and developing countries, the resources to each technology are misallocated in developing countries. For instance, entrepreneurs in the poorer country might not have enough access to capital or credit markets to raise the funds for the move to a newer and better technology. Or the older technology may be in the hands of older, elite groups with political power, who block access to new technologies that could spell their own ruin. Or local communities — again, possibly encouraged by domestic elites — might oppose new technologies for fear that these will let “foreign interests” into the country.

We will have much to say about these and related issues later in the book. But our quest for explanations starts with the limited input story. A developing country surely has less capital — physical and human — relative to labor, compared to its developed-country counterpart. Might this, and this alone, not explain persistent, ongoing differences in per capita income across countries? To understand such matters, we must expand our theory to allow capital to be systematically accumulated. That is the job of the fundamental “growth equation,” to which we now turn.

3.3.3 The Growth Equation. A little algebra at this stage will make our lives simpler. Divide time into periods $t = 0, 1, 2, 3, \dots$. We will tag relevant economic variables with the date: $Y(t)$ for output, $I(t)$ for investment, and so on. Investment augments the capital stock after accounting for depreciated capital, so in symbols:

$$K(t + 1) = (1 - \delta)K(t) + I(t),$$

where δ is the rate of depreciation. But we’ve just been going on about the famous macroeconomic balance condition, that investment equals savings. It follows that $I(t) = sY(t)$, where s is the *rate* of savings and $Y(t)$ is aggregate output, and using this in the equation above, we see that

$$(3.5) \quad K(t + 1) = (1 - \delta)K(t) + sY(t),$$

which tells us how the capital stock must change over time.

We’re going to convert this into per-capita terms by dividing by the total population, which we assume (only for expositional simplicity) to be equal to the active labor force $L(t)$. If we assume that population grows at a constant rate n , so that $L(t + 1) = (1 + n)L(t)$, (3.5) changes to

$$(3.6) \quad (1 + n)k(t + 1) = (1 - \delta)k(t) + sy(t),$$

where the lowercase k s and y s represent per capita magnitudes (K/L and Y/L , respectively).

This is our basic growth equation: make sure you understand the economic intuition underlying the algebra. It really is very simple. The right-hand side has two parts, depreciated per capita capital (which is $(1 - \delta)k(t)$) and current per capita savings (which is $sy(t)$). Added together, this should give us the new per capita capital stock $k(t + 1)$, except for one complication: population is growing, which exerts a downward drag on per capita capital stocks. This is why the left-hand side of (3.6) has the rate of growth of population (n) in it: the larger that rate, the lower is per capita capital stock in the next period.

Take the growth equation for a spin: start in your mind's eye with the per capita capital stock at any date t , $k(t)$. That produces per capita output $y(t)$ via the production function described in Section 3.3.2. Now the equation tells you what $k(t + 1)$ must be, and the story repeats, presumably without end.

3.4 The Solow Model

So far we've described the ingredients of a basic growth model, any growth model. A more detailed study must rely on a specific form of the production function, and we will start by presuming constant returns to scale and diminishing returns to each input, as in the Cobb-Douglas formulation of equation (3.2). The resulting analysis yields a fundamental and much-venerated theory of growth, due to Robert Solow (1957).

3.4.1 Dynamics of the Growth Equation. Figure 3.3 embeds a production function with diminishing returns into the growth equation (3.6). Mark any value of k on the horizontal axis, multiply the per capita output produced from it by s , which gives us fresh investment, and add the result to the depreciated value of k . The end product is the curved line in Figure 3.3, which looks very much like the production function itself (and indeed, is closely related), but has been transformed in the way we've just described. Figure 3.3 also plots the left-hand side of (3.6), the straight line $(1 + n)k$ as k changes. When there are diminishing returns, the curved line initially lies above this straight line and then falls below.⁵

⁵If you've been following the argument particularly closely, you will see that this last statement is exactly true if we make the additional assumption that the marginal product of capital is very high when there is very little capital and diminishes to zero as the per capita capital stock becomes very high.

Armed with this diagram, we can make some very strong predictions about growth. Figure 3.3 shows us two initial historical levels of the per capita capital stock — one “low” (Figure 3.3a) and one “high” (Figure 3.3b) — starting, in deference to the year I began writing this book, 1996. With the low stock, the marginal return to capital is very high and so the per capita capital stock expands quite rapidly. How do we see this in Figure 3.3? Well, we know from (3.6) that the supply of per capita capital is read off by traveling up to the point on the curved line corresponding to the initial stock $k(0)$. However, some of this supply is eroded by population growth. To find next year’s capital stock, travel horizontally until the line $(1+n)k$ is touched; the capital stock corresponding to this point is what you’re looking for. Now just iterate the process. We obtain the zigzag path in Figure 3.3a. Note that the growth of per capita capital *slows down* and that per capita capital finally settles close to k^* , which is a distinguished capital stock level where the curved and straight lines meet.

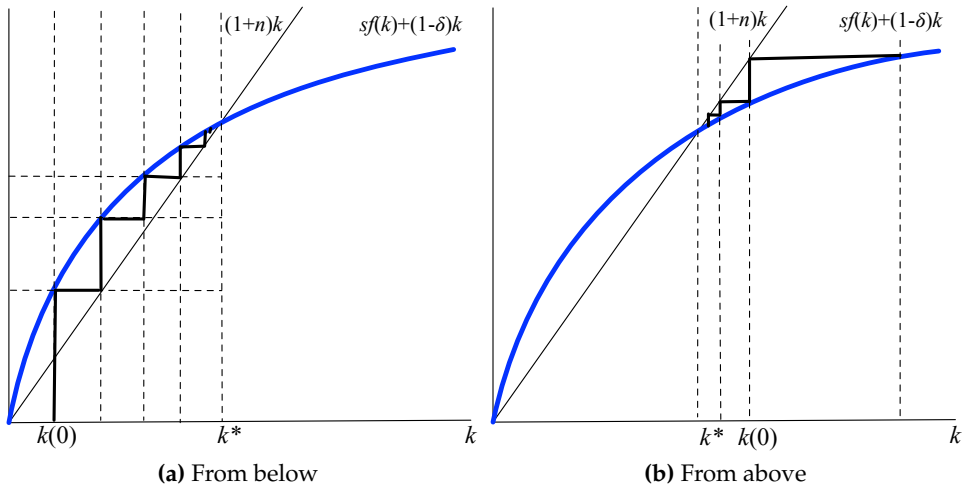


Figure 3.3. The Dynamics of the Solow Model.

Likewise, you may trace the argument for a high initial capital stock, as in Figure 3.3b. Here, there is an *erosion* of the per capita stock as time passes, with convergence occurring over time to the *same* per capita stock, k^* , as in Figure 3.3a. The idea here is exactly the opposite of that in the previous paragraph: the output-capital ratio is low, so the rate of expansion of aggregate capital is low. Therefore, population growth outstrips the rate of growth of capital, thus eroding the *per capita* capital stock.

3.4.2 The Steady State. We can think of k^* as a *steady-state* level of the per capita capital stock, to which the per capita capital stock, starting from *any* initial level, must converge.

In other words, growth in this model loses its momentum if capital is growing too fast relative to labor, which is precisely what happens to the left of k^* in Figure 3.3a. The reason is diminishing returns to capital, which lowers the marginal return to capital as it is accumulated faster than labor. That brings down the growth of capital in line with the growth of labor. This means that the long-run capital-labor ratio must be constant (and this is captured by the ratio k^*).

However, if the per capita capital stock settles down to some “steady-state” level, then so must per capita income! Thus in this version of the growth model, there is no long-run growth of per capita output, and *total* output grows precisely at the rate of growth of the population.

Indeed, *were* the economy to start from the steady-state level of k^* , it would stay at k^* in every period (after all, that is what the term “steady state” means). This means that in equation (3.6), we can put $k(t) = k(t + 1) = k^*$. If we use the notation y^* to denote the steady-state per capita output produced from k^* , and move terms around in (3.6) a bit, we obtain

$$(3.7) \quad \frac{k^*}{y^*} = \frac{s}{n + \delta}.$$

(You can connect k^* and y^* more explicitly with an equation like (3.3), if you’d like. Take a look at Section 3.4.5, where we do just this.)

At this point you must be scratching your head in some confusion. Or if you’re not, you should be. A growth model that predicts no growth? Do not fear, we will come back to this question.

3.4.3 How Parameters Affect the Steady State. The different parameters of the Solow model — the savings rate, population growth rate, or the rate of depreciation — do not affect the long-run growth rate of per capita income (which is zero). But they certainly affect the long-run *level* of income.

The effects of changes in various parameters can be easily studied using equation (3.7). An increase in s , the rate of savings, raises the right-hand side of the equation, necessitating an increase in the left-hand side to restore equality. This means that the new steady-state capital-output ratio — the value k^*/y^* — must be *higher*. With diminishing returns, that can only happen if the new steady-state level of y^* is higher as well. We thus see

that an increase in the savings rate raises the long-run level of per capita income. By exactly the same logic, we can check that an increase in the population growth rate or the rate of depreciation will lower the long-run level of per capita income.

But as always, make sure you understand the economics behind the algebra. For instance, a higher rate of depreciation means that more of national savings must go into the replacement of worn-out capital. This means that, all other things being equal, the economy accumulates a smaller net amount of per capita capital, and this lowers the steady-state level. You should similarly run through — verbally — the effects of changes in the savings rate and the population growth rate.

3.4.4 Technical Progress. We promised to address the no-growth paradox of the Solow model, and we shall do so at several points, starting now.

First a quick review of that counterfactual conclusion: given diminishing returns, the marginal contribution of capital to output must decline when capital grows faster than population, and increase when capital grows slower than population. That eventually forces the growth rate of aggregate capital — and output — to *equal* the growth rate of population, with zero growth per capita.

But this is the most parsimonious scaffolding of the model, which we can modify in several ways. For instance, the no-growth scenario disappears if there is continuing technical progress; that is, if the production function *shifts upward* over time as new knowledge is gained and applied. As long as the optimistic force of this shift outweighs the doom of diminishing returns, there is no reason why per capita growth cannot be sustained indefinitely.

This twist on the model makes intuitive sense. Imagine a world in which we never learnt to do things better. Then the influence of fixed factors — land, oil, and other resources — would make itself felt, and growth could conceivably vanish. In our simple model, we are benchmarking everything relative to labor by taking per-capita magnitudes, so labor plays the role of that fixed factor, and in the absence of technical progress, growth *per capita* must slow to a crawl.

A simple way to think about technical progress is that it contributes to the efficiency, or economic productivity, of labor. Let's make a distinction between the working population $L(t)$ and the amount of labor in "efficiency units" (call it $E(t)$). This distinction is necessary now because in the extension we're about to consider, the productivity of the working population

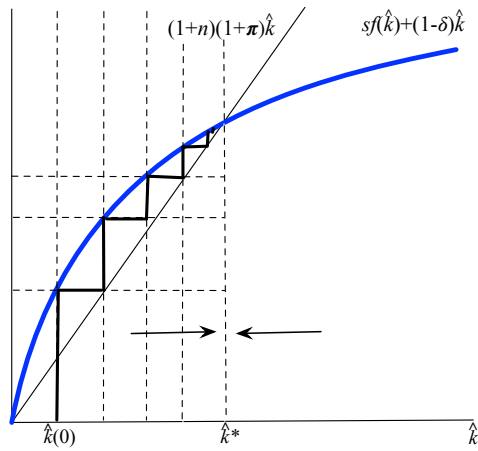


Figure 3.4. The Solow model with technical change.

is constantly increasing. The simplest way to capture this increase in productivity is to postulate that

$$(3.8) \quad E(t) = e(t)L(t),$$

where we can think of $e(t)$ as the productivity of an individual at time t . Not only does population grow over time (at the rate of n , just as before), but we now take it that efficiency grows too at the rate of π . Thus $e(t+1) = (1+\pi)e(t)$. We will refer to π as the rate of technical progress.

One more step and our adapted model is complete. Recall how we passed from (3.5) to (3.6) after dividing through by the working population to obtain per capita magnitudes. We do the same here, but with a twist: we divide by the *effective* population $e(t)L(t)$ to arrive at what looks like per capita capital and income, but these will really be magnitudes *per efficiency unit of labor*. Let's call them \hat{k} and \hat{y} to distinguish them from the earlier per capita values k and y . Performing the required division, we get something that looks very much like our old equation (3.6):

$$(3.9) \quad (1+n)(1+\pi)\hat{k}(t+1) = (1-\delta)\hat{k}(t) + s\hat{y}(t).$$

Now we simply apply our old ideas regarding production functions. Capital per efficiency unit of labor (\hat{k}) produces output per efficiency unit of labor (\hat{y}). As in the basic Solow model, if there is too much capital per efficiency unit, we have a shortage of (effective) labor and the output–capital ratio tends to fall: diminishing returns to an input still applies, but this time to *efficiency* units of labor. Thus Figure 3.4 repeats, in spirit, the analysis done for equation (3.6) in Figure 3.3. *Exactly* the same logic applies to equation (3.9).

Over time, the amount of capital per effective labor may rise or fall. Observe that if it is rising, this simply means that physical capital is growing faster than the rate of population growth and technical progress *combined*. However, by diminishing returns, output per efficiency unit rises less than proportionately. That tones down the subsequent growth rate of capital per efficiency unit. In Figure 3.4, this discussion corresponds to the region lying left of the intersection point \hat{k}^* . In this region, the growth rate of total capital falls over time as capital per efficiency unit rises. Likewise, to the right of the intersection, capital per efficiency unit falls over time. Once again, we have convergence, this time to the steady-state level \hat{k}^* .

To solve for the steady state, put $\hat{k}(t) = \hat{k}(t + 1) = \hat{k}^*$ in (3.9), just as we did to arrive at the steady state formula without technical progress; see (3.7). That gives us

$$(1 + n)(1 + \pi)\hat{k}^* = (1 - \delta)\hat{k}^* + s\hat{y}^*,$$

and rearranging this equation, we conclude that

$$\frac{\hat{k}^*}{\hat{y}^*} = \frac{s}{(1 + n)(1 + \pi) - (1 - \delta)}.$$

For an easy-to-remember version of this equation, note that both π and n are small numbers, such as 0.05 or 0.02, so their product is very small relative to the other terms and can be ignored in a first-order approximation. So multiplying out $(1 + n)(1 + \pi)$ and ignoring the term $n\pi$, we see that

$$(3.10) \quad \frac{\hat{k}^*}{\hat{y}^*} \simeq \frac{s}{n + \pi + \delta}.$$

(Once again, just as for equation (3.7), you can connect \hat{k}^* and \hat{y}^* with a specific production function, as we will do in Chapter 4, Section 4.3.3.)

So far the analysis runs parallel to the case of no technical progress. The novelty lies in the interpretation. *Even though* capital per efficiency unit converges to a stationary steady state, the amount of capital per member of the working population continues to increase. Indeed, the long-run increase in per capita income takes place precisely at the rate of technical progress.

In summary, think of two broad sources of growth: one is through better and more advanced methods of production (technical progress) and the second is via the continued buildup of plant, machinery, and other inputs that bring increased productive power.⁶ The Solow model claims that *in*

⁶This is not to deny that these two sources are often intimately linked: technical progress may be *embodied* in the new accumulation of capital inputs.

the absence of the first source, the second is not enough for sustained per capita growth. Viewed in this way, the Solow model is a pointer to studying the economics of technological progress, arguing that it is there that one must look for the ultimate sources of growth. This is not to say that such a claim is necessarily true, but it is certainly provocative and very far from being obviously wrong.

3.4.5 Another Route to Sustained Growth: The Harrod-Domar Model.

Let's take a second look at sustained growth. Eliminate technical progress for the moment, and return to the fundamental growth equation (3.6), reproduced here for easy reference:

$$(3.11) \quad (1 + n)k(t + 1) = (1 - \delta)k(t) + sy(t).$$

This turns into the Solow model under the *additional* assumption of diminishing returns to capital. That assumption creates variation in the capital-output ratio, allowing the economy to settle at the distinguished level k^* given by

$$(3.12) \quad \frac{k^*}{y^*} = \frac{s}{n + \delta'}$$

which is equation (3.7), again reproduced here. (This second equation is obtained, as already discussed, by setting $k(t) = k(t + 1) = k^*$ and $y(t) = y^*$ in the growth equation (3.11), and solving out for corresponding ratio k^*/y^* .) A solution exists because k/y varies from very high to very low values over the range of the production function.⁷ For instance, when the production function is Cobb-Douglas, so that

$$y = Ak^\alpha,$$

the steady state condition (3.12) becomes

$$\frac{k^*}{Ak^{*\alpha}} = \frac{s}{n + \delta'}$$

so that after a little not-so-tedious algebra, we see that

$$(3.13) \quad k^* = \left(\frac{sA}{n + \delta'} \right)^{1/(1-\alpha)}.$$

This is a simple and rewarding equation, because you can “see” the steady state as explicitly as possible, and what is more, you can redo all the parametric exercises of Section 3.4.3 in a flash simply by eyeballing (3.13).

But there is something else that I'd like to draw your attention to. Look at the “diminishing returns parameter” α and remember that the smaller

⁷The careful reader will observe again that we're assuming diminishing returns plus suitable end-point conditions on the marginal product of capital.

it is, the more it is that returns diminish, while at the other end, as α becomes close to 1, the production function becomes almost linear and exhibits constant returns to capital. As we bring α up close to 1, the steady state level of capital becomes ever larger (if $sA > n + \delta$) or ever smaller (if $sA < n + \delta$) and at α equal to 1, when the production function is *exactly* linear in capital, there is no steady state: the economy either grows to infinity or shrinks to zero! This is not an algebraic sleight of hand; on the contrary, it makes intuitive sense. The easiest way to see it is to appreciate that when $\alpha = 1$, the current scale of the economy (proxied by k) is irrelevant: whatever rate it can grow at k , it can replicate that at $2k$, $3k$, or a million times k . With diminishing returns, that isn't possible, which is why a steady state is ultimately reached, but when there is constant returns to the capital input, everything can grow or decline at exactly the same rate, irrespective of scale.

The fundamental growth equation (3.11) can handle this without a problem, provided we don't go down the garden path looking for steady states where there are none to be found. With constant returns to scale, the output-capital ratio is a constant; in fact, it is exactly A in the Cobb-Douglas production function with $\alpha = 1$: $y = Ak$. Using this in the growth equation yields

$$(1 + n)k(t + 1) = (1 - \delta)k(t) + sAk(t),$$

and after the inevitable moving-around of terms, we have

$$(3.14) \quad \text{Rate of growth} = \frac{k(t + 1) - k(t)}{k(t)} = \frac{sA - (n + \delta)}{1 + n}.$$

This is an influential relationship, known as the *Harrod-Domar equation*, named after Roy Harrod and Evsey Domar, who wrote some of the earliest papers on the subject in 1939 and 1946, respectively. The left-hand side is just the rate of growth of the per capita capital stock, and by the linearity of the production function it is also the rate of growth of per capita *income*. Time appears here on the left-hand side, but it doesn't in the rest of the equation, which shows that the rate of growth is constant and unchanging. What is more, it provides a formula for that rate of growth, given by the right-hand side of (3.14).

For an easy-to-remember version of (3.14), let g stand for the rate of growth, multiply through by $1 + n$, and note that both g and n are small numbers (analogous to our approach to deriving (3.10)), and their product gn is therefore *extra* small relative to these numbers. That gives us the approximation

$$(3.15) \quad sA \simeq g + n + \delta,$$

which can be used in place of (3.14) without much loss of accuracy.

It isn't hard to see why this equation was influential. It has the air of a recipe. Thomas Piketty, in his book *Capital* (2014) — on which more later — calls it the “second fundamental law of capitalism.” The equation firmly links the growth rate of the economy to fundamental variables, such as the ability of the economy to save, the capital–output ratio and the growth rate of population. And capitalism apart, central planning in countries such as India and the erstwhile Soviet Union was deeply influenced by the Harrod–Domar equation (see box).

Growth Engineering: The Soviet Experience

The Harrod–Domar model, as we have seen, has both descriptive and prescriptive value. The growth rate depends on certain parameters and, in a free market economy, these parameters are determined by people's tastes and technology. However, in a socialist, centrally planned economy (or even in a mixed economy with a large public sector), the government may have enough instruments (such as direct control over production and allocation, strong powers of taxation and confiscation, etc.) to manipulate these parameters to influence the growth rate. Given a government's growth objectives and existing technological conditions (e.g., the capital–output ratio), the Harrod–Domar model can be used to obtain policy clues; for example, the desired rate of investment to be undertaken so as to achieve this aim.

The first controlled experiment in “growth engineering” undertaken in the world was in the former Soviet Union, after the Bolshevik Revolution in 1917. The years immediately following the Revolution were spent in a bitter struggle—between the Bolsheviks and their various enemies, particularly the White Army of the previous Czarist regime—over the control of territory and productive assets such as land, factories, and machinery. Through the decade of the 1920s, the Bolsheviks gradually extended control over most of the Soviet Union (consisting of Russia, Ukraine, and other smaller states) that encompassed almost the whole of industry, channels of trade and commerce, food-grain distribution, and currency. The time had come, therefore, to use this newly acquired control over the economic machinery to achieve the economic goals of the revolutionary Bolsheviks, the foremost among these goals being a fast pace of industrialization.⁸

⁸On the eve of the Revolution, Russia took a back seat among European nations in extent of industrialization, despite a rich endowment of natural resources. According to the calculations of P. Bairoch, based on the per capita consumption of essential industrial inputs, namely, raw cotton, pig iron, railway services, coal, and steam power, Russia ranked last among nine major European nations in 1910, behind even Spain and Italy. See Nove [1969].

	1927–28 (<i>actual</i>)	1932–33 (<i>plan</i>)	1932 (<i>actual</i>)
National income	24.4	49.7	45.5
Gross industrial production	18.3	43.2	43.3
(a) Producers' goods	6.0	18.1	23.1
(b) Consumers' goods	12.3	25.1	20.2
Gross agricultural production	13.1	25.8	16.6

Source: Dobb [1966].

^aAll figures are in 100 million 1926–27 rubles.

Table 3.3. Targets and achievements of the first Soviet five Year Plan (1928–29 to 1932–33).^a

Toward the end of the 1920s, the need for a coordinated approach to tackle the problem of industrialization on all fronts was strongly felt. Under the auspices of the State Economic Planning Commission (called the *Gosplan*), a series of draft plans was drawn up which culminated in the first Soviet Five Year Plan (a predecessor to many more), which covered the period from 1929 to 1933. At the level of objectives, the plan placed a strong emphasis on industrial growth. The resulting need to step up the rate of investment was reflected in the plan target of increasing it from the existing level of 19.9% of national income in 1927–28 to 33.6% by 1932–33. (Dobb [1966, p. 236]).

How did the Soviet economy perform under the first Five Year Plan? Table 3.3 shows some of the plan targets and actual achievements, and what emerges is quite impressive. Within a space of five years, real national income nearly doubled, although it stayed slightly below the plan target. Progress on the industrial front was truly spectacular: gross industrial production increased almost 2.5 times. This was mainly due to rapid expansion in the machine producing sector (where the increment was a factor of nearly 4, far in excess of even plan targets), which is understandable, given the enormous emphasis on heavy industry in order to expand Russia's meager industrial base. Note that the production of consumer goods fell way below plan targets.

An equally spectacular failure shows up in the agricultural sector, in which actual production in 1932 was barely two-thirds of the plan target and only slightly more than the 1927–28 level. The reason was probably that the Bolsheviks' control over agriculture was never as complete as that over industry: continuing strife with the *kulak* farmers (large landowners from the Czarist era) took its toll on crop production.

Unlike the Solow model, the Harrod-Domar variant it generates sustained growth *without having to invoke technical progress*. And yet, does that make

it a “better” model? Not necessarily. There is a price to be paid for the result, which is that the model assumes that labor is not a serious constraint on production at all (if it were, the returns to ever more capital, holding labor constant, would presumably diminish). But matters are a bit more complicated than that. In real life, there aren’t just two inputs, “capital” and “labor.” There is a whole host of them, and the predictive power of Harrod-Domar versus Solow would ultimately stand or fall on the *combined* diminishing returns — or lack thereof — of *all* those inputs relative to raw labor.

We will be back to Solow versus Harrod-Domar at a later stage. For now, my only request is that you don’t get confused. It doesn’t matter that one model generates sustained growth while the other doesn’t. They rely, as I’ve tried to make clear, on different assumptions. Therefore, which model is more relevant is ultimately an empirical question and, as we shall soon see, the jury is still out on the issue. For now, the takeaway point is that constant technical progress is needed to generate sustained growth in the face of diminishing returns to capital, while capital accumulation per se is enough to do the trick when the return to capital is constant (or increasing).