Chapter 3

The Solow model

One of the long-run economic facts presented in Chapter 2 is the stability of the capital-output ratio K_t/Y_t (where Y_t represents aggregate output at period t and K_t represents aggregate capital at period t) over time: capital is roughly three times annual GDP. Earlier contributions of growth theory, such as the so-called Harrod-Domar model, considered this stability as a technological property and incorporated it as one of the assumptions in the model. The capital stock, traditionally divided into structures and equipment but nowadays also containing some intangible components (e.g., software), is one of the important production inputs, but it is of course not the only one. It is possible to produce products in a highly capital-intensive way, but clearly there is a choice and using labor—different people's time, skills, and effort—is the most obvious alternative, or complement. Given the many possibilities in which production process can be organized, it is therefore not obvious why, at the macro level, K_t/Y_t is almost constant over time. As argued in Chapter 2, this was Solow's starting point. He reconciled the stable capital–output ratio with the substitutability of capital and labor through a sequence of insights that gave rise both to a framework for analyzing macroeconomic dynamics and to systematic methods for measuring technological change. The purpose of this chapter is to present the Solow model, outlining its key assumptions and their implications.

A central element in the Solow model is the aggregate production function. In the aggregate production function, there are three economic variables that can affect the growth of GDP: technology A_t, capital K_t, and labor L_t. The Solow model focuses on the endogenous accumulation of capital K_t. We will see that K_t not only reacts to the saving rate but also to A_t and L_t. After solving the model, which will deliver a stable capital-output ratio, we will focus on two main takeaways: (i) the fundamental source of long-run growth in per capita income is the growth in A_t; (ii) if all parameter values are common, different economies converge to the same (both in terms of level and growth rate) income per capita in the long run.

3.1 The basic model

We start our exposition using the simplest version of the model, where there is neither technological progress nor any growth in the size of the population or the skills of workers.

The centerpiece of the Solow model is the aggregate production function

$$Y_t = F(K_t, L_t).$$

Note that we can interpret Y_t as the GDP in this economy. We make the following assumptions for the function $F(K, L)$:

1. $F(K, L)$ is strictly increasing in both K and L.

2. $F(K, L)$ is strictly quasiconcave in (K, L) (it has strictly convex isoquants).

3. $F(K, L)$ exhibits constant returns to scale in (K, L): when K and L change to cK and cL, with any $c > 0$, $F(K, L)$ becomes $cF(K, L)$.

4. $F(0, L) = 0$.

5. $\lim_{K \to 0} F_1(K, L) = \infty$, where $F_1(K, L) \equiv \partial F(K, L)/\partial K$.

6. $\lim_{K \to \infty} F_1(K, L) = 0$.

Assumptions 5 and 6 are often called Inada conditions and are stronger than we need but these assumptions simplify the exposition.[1]

In the basic model, we assume that the population is constant and that hours worked per worker is constant, so that L_t is constant. We normalize both population and hours per capita to 1; therefore, the only variable input for production is K_t, and because of this normalization, we can write $F(K_t, 1) = F(k_t, 1)$, where we remind the reader that lower-case letters are per-capita measures. Let us use $f(k_t)$ to denote $F(k_t, 1)$. Then the production function can be expressed

$$y_t = f(k_t). \tag{3.1}$$

The second important piece of the Solow model is the equation that describes the evolution of the capital stock:

$$k_{t+1} = i_t + (1 - \delta)k_t, \tag{3.2}$$

where i_t is investment in period t. The existing capital stock loses value, δk_t, while being used, where $\delta \in (0, 1)$ is capital's depreciation rate.

These two centerpieces are connected through individual behavior. First, the goods supply y_t is equal to the demand for goods, $c_t + i_t$:

$$y_t = c_t + i_t. \tag{3.3}$$

Here, c_t is consumption and i_t is investment (both per capita). Note that we implicitly assume (as in the large part of the following chapters) that goods are homogeneous and can be used for both consumption and investment. Because output y_t is also the total income for consumers, it is either consumed or saved. Therefore, we know that total saving has to equal total investment. In an open economy—where there is trade—this does not necessarily

[1] We also assume that F is twice continuously differentiable. This means that first-order conditions to maximization problems involving F can be differentiated and then generate continuous functions.

hold. Finally, in this chapter section, we can interpret both c and i as including government consumption and investment, respectively.

In the Solow model, instead of explicitly modeling the consumption-saving decisions of consumers, the consumers are assumed to mechanically save a constant fraction of their income, so that investment is given by

$$i_t = sy_t, \tag{3.4}$$

where $s \in (0, 1)$ is the constant saving rate. This behavioral assumption is relaxed and replaced by consumers' optimizing consumption-saving behavior in Chapter 4.

Inserting equation (3.4) into equation (3.2) and using equation (3.1) yields

$$k_{t+1} = (1 - \delta)k_t + sf(k_t). \tag{3.5}$$

This difference equation expresses the dynamics of the capital stock k_t over time. This is *the fundamental equation of the Solow model*. Note that the only endogenous variable on the right-hand side of the fundamental equation is k_t. Therefore, the next period's stock of capital k_{t+1} can be determined only with the knowledge of the current capital stock k_t, given values of exogenous objects: the scalars δ and s and the function f. Note also that, starting from a given k_0, once we obtain the series of $\{k_{t+1}\}_{t=0}^{\infty}$ from the fundamental equation (3.5), the time series $y_t = f(k_t)$, $c_t = (1 - s)y_t$, and $i_t = sy_t$ can readily be obtained.

3.1.1 Steady state and dynamics

To analyze the difference equation (3.5), we first consider a special situation where k_t is constant over time. Call this situation the *steady state* and denote it with an upper bar: $k_t = \bar{k}$ for all t. From the fundamental equation (3.5), the steady-state capital stock can be determined by the solution of the equation

$$\bar{k} = (1 - \delta)\bar{k} + sf(\bar{k}).$$

This equation implies $\delta\bar{k} = sf(\bar{k})$. It is straightforward to verify that, under the assumptions for the aggregate production function in the previous section, a strictly positive value of $\bar{k}$ that solves this equation always exists and is unique. We can analyze the steady state of the model by solving the equation graphically. Plot the left-hand side, δk, and the right-hand side, $sf(k)$ as functions of k. The left-hand side is a straight line through the origin with a positive slope. The right-hand side is strictly increasing and strictly concave, reflecting diminishing returns to capital. The slope of the right-hand-side is infinity at the origin and approaches 0 as $\bar{k} \to \infty$. Clearly, we see that an intersection exists and is unique. The Inada conditions are used here to guarantee the existence of $\bar{k}$. In the context of the basic model and as pointed out above, the Inada conditions are stronger than necessary: they can be replaced by weaker versions $\lim_{k \to 0} F_1(k, 1) > \delta/s$ and $\lim_{k \to \infty} F_1(k, 1) < \delta/s$.

A particularly useful production function is the Cobb-Douglas production function: $F(K, L) = K^{\alpha}L^{1-\alpha}$, so that $f(k) = k^{\alpha}$, where $\alpha \in (0, 1)$. This production function satisfies all assumptions we need, including the Inada conditions. With Cobb-Douglas production, $\bar{k}$ can be solved for analytically:

$$\bar{k} = \left(\frac{s}{\delta}\right)^{\frac{1}{1-\alpha}}. \tag{3.6}$$

From this expression, we can see that the capital stock in the steady state is increasing in the saving rate s and decreasing in the depreciation rate δ. Because aggregate output (GDP) is $\bar{y} = f(\bar{k})$, $\bar{y}$ is also increasing in s and decreasing in δ.

Now, let us use a diagram to analyze the dynamics of k_t when $k_0 > 0$ is not at the steady-state level. Figure 3.1 plots the equation (3.5) with the 45-degree line (that is, representing $k_{t+1} = k_t$). In the figure, the intersection of (3.5) and the 45-degree line represents the steady-state $\bar{k}$.

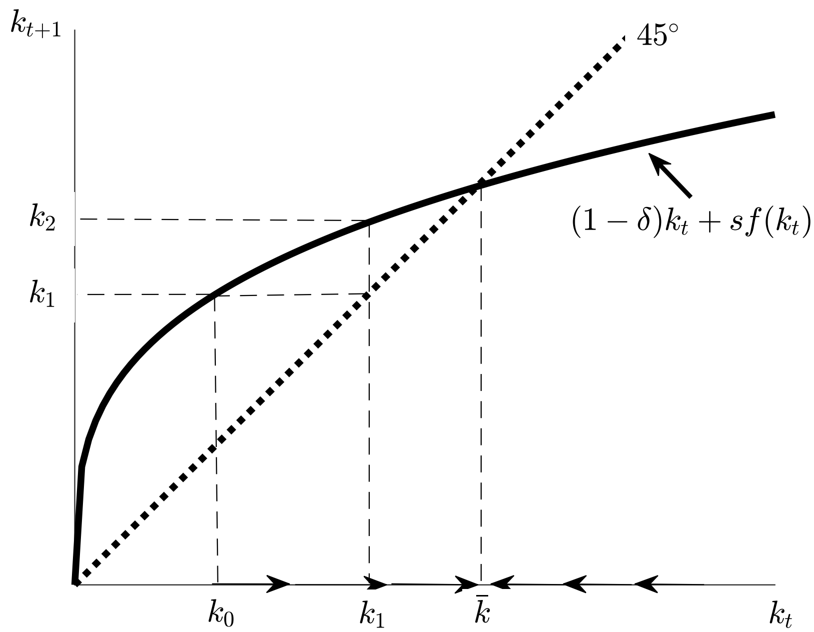

Figure 3.1: Dynamics in the Solow model.

In the figure, when we start from a given k_0 on the horizontal axis, we can obtain k_1 on the vertical axis by using the (3.5) curve. By placing this k_1 back on the horizontal axis and using the (3.5) curve again, we obtain k_2, and so on. This procedure yields the full time series of the capital stock: $\{k_{t+1}\}_{t=0}^{\infty}$. One can easily verify the dynamics of k_t exhibits a global and monotonic convergence to $\bar{k}$, regardless of the initial value k_0. That is, whatever the starting point is, the time series of k_t gradually approaches $\bar{k}$ over time.

The figure provides useful intuition, but how would a mathematical proof be put together? Suppose that $k_t < \bar{k}$. It is then straightforward to show that (i) $k_{t+1} > k_t$ (because $sf(k_t) > \delta k_t$ when $k_t < \bar{k}$) and also that (ii) $k_{t+1} < \bar{k}$ (because $k_t < \bar{k}$ and the right-hand side of (3.5) is increasing in k_t). Repeating this procedure, we can see the sequence $\{k_t, k_{t+1}, k_{t+2}, ...\}$ is monotone and bounded by $[k_t, \bar{k}]$. From the Monotone Convergence Theorem, the sequence has a limit. The limit has to be $\bar{k}$, as the limit is unique under the conditions given.

Intuitively, the convergence occurs because the aggregate production function $f(k_t)$ has decreasing returns to capital (Assumption 2 above). Equation (3.2), rewritten in terms of the change in the capital stock,

$$k_{t+1} - k_t = i_t - \delta k_t,$$

reveals two forces that go in opposite directions: (gross) investment and depreciation. When the total capital stock is small, output per unit of capital is large, and the constant saving rate then implies that a large (gross) investment is made relative to the existing capital stock. This process enables the aggregate capital stock to increase. As k_t increases, output per unit of capital becomes smaller due to the decreasing returns property, and when k_t is very large enough, the gross investment cannot cover total deprecation, δk_t. Thus, the investment force is stronger when k_t is small and the depreciation force is stronger when k_t is large. This relationship is perhaps even clearer if we write the above equation in terms of the growth rate:

$$\frac{k_{t+1} - k_t}{k_t} = \frac{sf(k_t)}{k_t} - \delta,$$

where we have replaced $i_t = sf(k_t)$. The assumptions $f''(k_t) < 0$ and $f(0) = 0$ imply that $f(k_t)/k_t$ is decreasing in k_t, generating the negative relationship between the investment force of pushing up the capital stock and the level of the capital stock. Indeed, when saving behavior (represented here by s) is modeled as an explicit choice, it can offset the forces driving convergence—but not strongly enough to overturn the convergence result. This issue will be discussed in detail in the next chapter.

Let us go back to our motivating fact: the stability of k_t/y_t over time. In the basic model here, k_t/y_t is of course constant in steady state, as are all the variables. However, as we will see below, even in a growing economy where both k_t and y_t increase over time, the ratio k_t/y_t eventually stabilizes and becomes a constant.

In the Cobb-Douglas case above, the steady-state k/y ratio can be solved out as

$$\frac{\bar{k}}{\bar{y}} = \frac{s}{\delta}.$$

Clearly, in the long run k_t/y_t is larger when s is larger and when δ is smaller.

Other kinds of dynamics

The growth model can, in principle, generate very rich (and complex!) dynamics if its neoclassical feature is not present, i.e., if the production function is not strictly concave in capital. There are applications in the economics literature that, in reduced form, have such non-neoclassical features, and we now briefly illustrate how they can work.

Endogenous growth Consider the situation where $F_1(k, 1)$ is uniformly above δ/s. Figure 3.2 below draws such an example. In this case, the steady state with $\bar{k} > 0$ does not exist, and k_t keeps growing larger over time. That is, there is unbounded growth "by itself": growth in *endogenous*. This concept will be discussed more in Chapter 13 below but in a richer model where other production inputs can also be accumulated. Here, given that one expects decreasing returns to each input—such as capital—it is hard to take this case very seriously.

In the special (and illustrative) case where F_1 is a constant—as illustrated in the graph—we can think of output as linear in capital: $y_t = Ak_t$, with no role for labor (make $\alpha = 1$ in the Cobb-Douglas setting).[2] Given that labor commands about two thirds of the income

[2] One can imagine a role for labor if $Y_t = AK_t + BL$, which is CRS, but here labor would not matter asymptotically if $A > \delta/s$.

from production, this setup does not seem empirically plausible. In a setup with endogenous growth such as this, two identical countries starting out with different capital stocks will be forever different; the gap between them, in percentage terms, will stay constant.

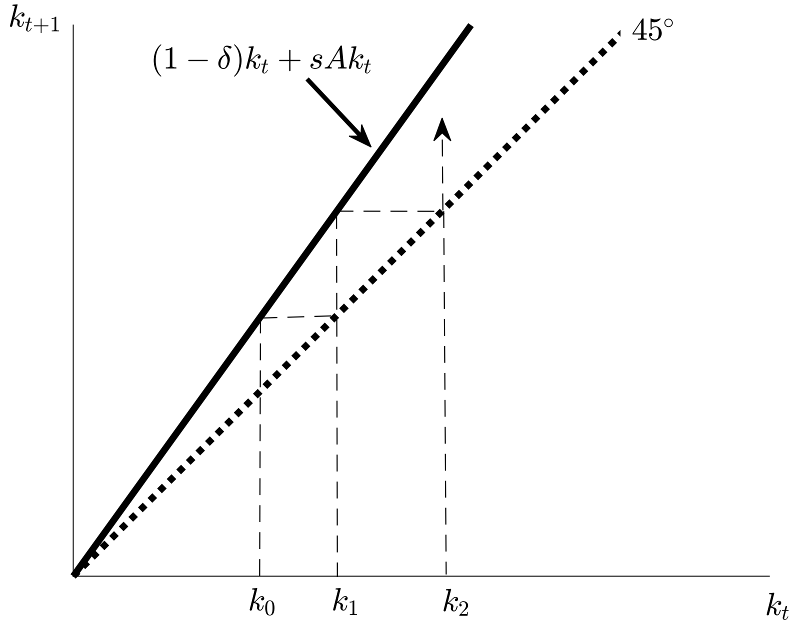

Figure 3.2: Endogenous growth in the Solow model.

Poverty traps

Suppose that the production function is not globally concave in k: it has a middle section that is convex. This could be true if there are some regions of k with increasing returns, say, as a result of large infrastructure investments—the building of transportation networks. In such a case, the right-hand side of (3.5) will not be concave, and it may cross the 45-degree line multiple times, as illustrated in Figure 3.3 below.

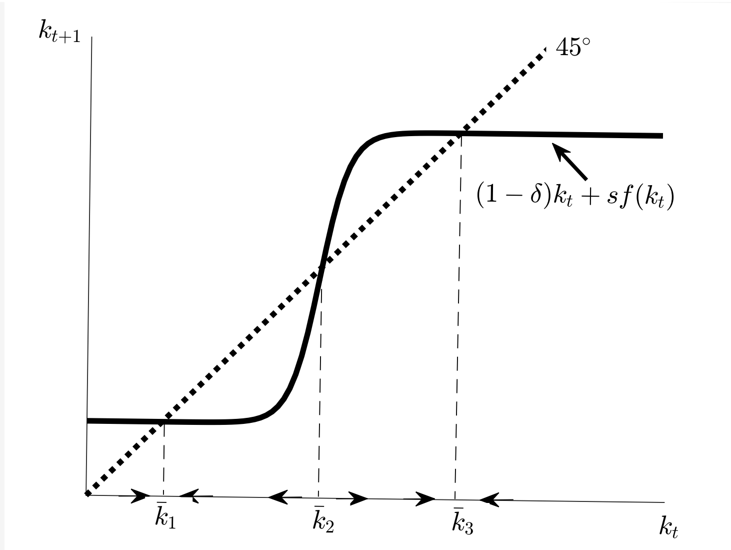

Figure 3.3: Poverty traps in the Solow model

Clearly, in this case there are multiple steady states and at least one of the steady states will then not be "stable": k_t will not converge to that steady state even when k_0 starts very close to it (a small perturbation away from the steady state leads further away from it). When there are multiple steady states, an economy can get stuck in the steady state with a low $\bar{k}$ (and a low GDP) when it starts from a low k_0. This situation is often called the *poverty trap*. The gap between a poor country with a small k_0 and a rich country with a large k_0 may never close in this setting. In Figure 3.3, there are three steady states, $\bar{k}_1$, $\bar{k}_2$, and $\bar{k}_3$. Of these three, $\bar{k}_1$ and $\bar{k}_3$ are stable steady states. When the economy starts from a very low k_0, the economy converges to $\bar{k}_1$ and gets trapped in it. To escape from the trap, the capital stock would have to be pushed up to a larger level than $\bar{k}_2$, from which it would converge to $\bar{k}_3$. One way to achieve this movement is to (temporarily) encourage very high saving. If the saving rate is raised sufficiently such that the $(1-\delta)k_t + sf(k_t)$ curve moves up sufficiently, the steady states $\bar{k}_1$ and $\bar{k}_2$ will disappear and the economy converges globally to $\bar{k}_3$. The growth/development literature does not appear to have identified sufficiently large increasing returns leading to results of the kind described here, but it is an interesting possibility.

Non-monotonic dynamics and chaos An even more radical departure from the neoclassical setting is if $f(k)$ declines in k at high levels of k. Conceptually, if a bakery has no ovens, ovens have high marginal productivity, and as more ovens are added, the marginal productivity declines, and it will become negative once there are so many ovens in the bakery that there is neither space for bakers nor for the dough. This possibility is more esoteric in a macroeconomic context but let us nevertheless study it briefly. So when $f(k)$ decreases steeply enough, the right-hand side of (3.5) will become decreasing in k_t. Illustrating this graphically, we will see that convergence, if convergence is at all possible, will not be monotonic.[a] In fact, k_t can exhibit forever oscillating (or even chaotic) dynamics.[b] An example of is drawn in Figure 3.4.

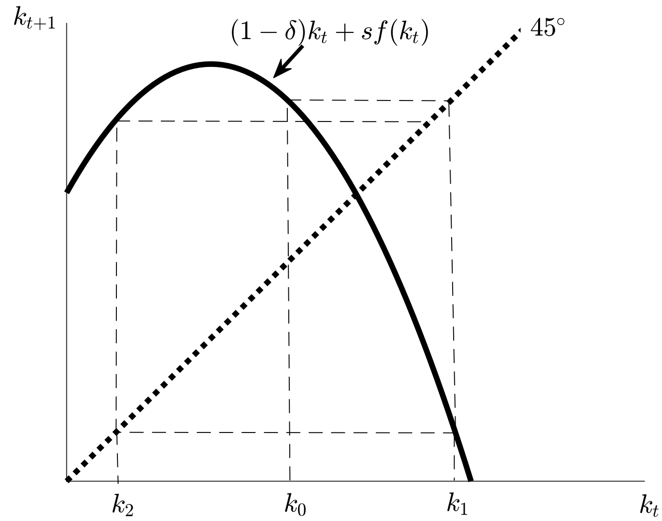

Figure 3.4: Complex dynamics in the Solow model

[a]Locally stable dynamics will occur if the slope at the steady state is less than 1 in absolute value.

[b]Chaos is a mathematical term; it involves great sensitivity to initial conditions and forever non-monotone behavior that never settles down to a repeated pattern (a repeated pattern could be a two-cycle): it looks "random."

3.2 The growing economy

Now we extend the model to the situation where A_t and L_t grow over time. In the previous section, the long-run outcome was the steady state where there is no growth. This extension is necessary for addressing the facts related to the growth issues described in Chapter 2.

We assume that the aggregate production function takes the form of

$$Y_t = F(K_t, A_t L_t).$$

There are two changes from the basic model: first, we allow the labor input (population times hours worked per person) L_t to grow over time. Second, and more importantly, we allow for technological progress. In this production function, the variable representing the technology level, A_t, is multiplied by labor input L_t. Technological progress takes a form of improving the labor input, and $A_t L_t$ is often referred to as the total number of *efficiency units of labor* (or *effective labor*). This form of technological progress is labor-augmenting; it was introduced in the previous chapter.[3] As was also asserted there, Uzawa (1961) proved that labor-augmenting technical change is the only form of technical progress that is consistent with exact balanced growth, that is, the growth path where aggregate variables such as output and capital grow at a constant rate. Uzawa's theorem is formally stated and proved in Appendix 3.A.

We assume that the (net) growth rate of A_t is γ and that the growth rate of L_t is n.[4]

[3]This form of technical progress is sometimes also called Harrod-neutral.

[4]n has two origins: a growing population and changes in hours worked per person, which we saw from Chapter 2 is best characterized by a decline in the longer run.

The same manipulations of equations as in the basic model yield

$$K_{t+1} = (1-\delta)K_t + sF(K_t, A_t L_t).$$

Dividing both sides by $A_t L_t$, we obtain

$$\frac{K_{t+1}}{A_t L_t} = (1-\delta)\frac{K_t}{A_t L_t} + s\frac{F(K_t, A_t L_t)}{A_t L_t}.$$

Let us define a new variable $\tilde{k}_t$ by

$$\tilde{k}_t \equiv \frac{K_t}{A_t L_t}.$$

Then, because the above equation can be rewritten as

$$\frac{A_{t+1}}{A_t}\frac{L_{t+1}}{L_t}\frac{K_{t+1}}{A_{t+1}L_{t+1}} = (1-\delta)\frac{K_t}{A_t L_t} + sF\left(\frac{K_t}{A_t L_t}, 1\right),$$

we obtain

$$(1+\gamma)(1+n)\tilde{k}_{t+1} = (1-\delta)\tilde{k}_t + sf(\tilde{k}_t). \tag{3.7}$$

Here, $f(\tilde{k}_t) \equiv F(\tilde{k}_t, 1)$ as in the basic model. After the capital stock K_t is normalized by $A_t L_t$, we obtain a very similar difference equation as the fundamental equation (3.5) in the basic model. Once we characterize the dynamics of $\tilde{k}_t$, we can "untransform" it into the core macro variables Y_t, K_t, and C_t.

3.2.1 Balanced growth and dynamics

The characterization of the fundamental equation (3.7) follows similar steps as for the basic model. The concept corresponding to the steady state in the basic model is the *balanced growth path* (some researchers still prefer to use the name "steady state" for the balanced growth path, because the normalized variables are "steady" also in this case). Along the balanced growth path, the normalized capital stock $\tilde{k}_t$ is constant, and typical economic variables, such as Y_t, K_t, and C_t, grow at a constant rate. Once again, we use a notation with upper bar: $\tilde{k}_{t+1} = \tilde{k}_t = \bar{\tilde{k}}$. The value of $\bar{\tilde{k}}$ along the balanced-growth path solves

$$(1+\gamma)(1+n)\bar{\tilde{k}} = (1-\delta)\bar{\tilde{k}} + sf(\bar{\tilde{k}}).$$

With a Cobb-Douglas production function we can obtain a closed-form solution:

$$\bar{\tilde{k}} = \left(\frac{s}{(1+\gamma)(1+n)+\delta-1}\right)^{\frac{1}{1-\alpha}}. \tag{3.8}$$

The dynamic property of the model can be analyzed similarly to the basic model. As in the basic model, starting from any $\tilde{k}_0$, $\left\{\tilde{k}_{t+1}\right\}_{t=0}^{\infty}$ converges monotonically to $\bar{\tilde{k}}$. In (3.8), the rate of technological progress γ and the population growth rate n work similarly to depreciation: maintaining a level of $k_t = K_t/(A_t L_t)$ is harder as A and L grow faster; i.e., each unit of untransformed capital needs to grow faster, as if making up for depreciation.

In this framework, we can analyze how various economic variables grow over time. For example, suppose L_t is simply population size (assuming that all citizens work one unit) and then consider income per capita, defined as $y_t \equiv Y_t/L_t$. Because $Y_t/(A_tL_t) = f(\tilde{k}_t)$, in the long run, $Y_t/(A_tL_t)$ converges to $f(\bar{\tilde{k}})$. Therefore, in the long run, income per capita converges to

$$y_t = f(\bar{\tilde{k}})A_t$$

and the growth rate of y_t in the long run is

$$\frac{y_{t+1} - y_t}{y_t} = \frac{f(\bar{\tilde{k}})A_{t+1} - f(\bar{\tilde{k}})A_t}{f(\bar{\tilde{k}})A_t} = \frac{A_{t+1} - A_t}{A_t} = \gamma.$$

The growth in technology A_t is essential in sustaining long-run growth in per capita income. Surprisingly, no other parameters affect the long-run growth of per capita income. For example, encouraging saving (an increase in s) does not affect the long-run growth rate of per capita income in the economy. Note that this result does not mean that the change in s does not have any effect on economic outcome: it has an effect on the *level* of per capita income, rather than the growth rate. It also has an effect on the growth rate in the short run (when the economy is not yet on the balanced-growth path).

In the short run, $\tilde{k}_t$ changes over time and its movement has an effect on the economic outcome. For example, the growth rate per capita income is now

$$\frac{y_{t+1} - y_t}{y_t} = \frac{f(\tilde{k}_{t+1})A_{t+1} - f(\tilde{k}_t)A_t}{f(\tilde{k}_t)A_t} = \frac{f(\tilde{k}_{t+1})}{f(\tilde{k}_t)}(1+\gamma) - 1.$$

From the fundamental equation (3.7), we know that when $\tilde{k}_t < \bar{\tilde{k}}$, $\tilde{k}_t$ increases over time, that is, $\tilde{k}_{t+1} > \tilde{k}_t$. Therefore, in this case, $f(\tilde{k}_{t+1})/f(\tilde{k}_t) > 1$ and the growth rate of y_t in the short run is larger than γ. Similarly, when $\tilde{k}_t > \bar{\tilde{k}}$, the growth rate of y_t is smaller than γ. In other words, the Solow model predicts that income per capita of a poor country grows faster than at rate γ and that income per capita of a rich country grows slower than at rate γ in the short run. This difference in growth rates is another representation of the convergence prediction of the Solow model.

3.3 Stylized facts and the Solow model

The model with growth, presented above, can match various stylized facts of economic growth. First, going back to our motivating fact on K_t/Y_t, because $Y_t/(A_tL_t) = f(\tilde{k}_t)$ and $K_t/(A_tL_t) = \tilde{k}_t$ are both constant along the balanced growth path, $K_t/Y_t = \tilde{k}_t/f(\tilde{k}_t)$ is also constant in the long run. Once again, the Solow model can replicate the constant K_t/Y_t in the data.

The first fact in Chapter 2 was the steady growth of the GDP per capita. As we have seen above, the GDP per capita grows at the rate γ along the balanced growth path (towards which the economy converges from any starting point). This fact, therefore, is consistent with the Solow model with technological progress.

Another stylized fact is that the return to physical capital has been nearly constant. Here, we need to first compute the return to physical capital in the model. Suppose that firms maximize profit under competitive markets:

$$\max_{K_t, A_t L_t} F(K_t, A_t L_t) - r_t K_t - w_t A_t L_t. \tag{3.9}$$

Here, output is taken as the numéraire and its price is set at one. Therefore, $F(K_t, A_t L_t)$ is the revenue and $r_t K_t + w_t A_t L_t$ is the cost. Let us assume, for simplicity, that the capital stock is owned by the consumers and rented to the firms with the rental rate r_t. Thus, r_t represents the return to capital. The other component of the cost is the wage payment: w_t is the wage per efficiency unit of labor. From the first-order condition for K_t, the return to physical capital is equal to the marginal product of capital:

$$r_t = F_1(K_t, A_t L_t).$$

Differentiating both sides of the equation $f(K_t/A_t L_t) = F(K_t, A_t L_t)/(A_t L_t)$ with respect to K_t, we obtain that

$$r_t = f'(\tilde{k}_t).$$

Along the balanced growth path, therefore, r_t is constant because the right-hand side is constant at $f'(\bar{\tilde{k}})$.

Another prominent fact is the stability of the labor share and the capital share. The capital share is equal to $r_t K_t / Y_t$, and it can readily be seen that it is constant, because we have already seen that both r_t and K_t/Y_t are constant along the balanced-growth path. The labor share is $w_t A_t L_t / Y_t$. From the first-order condition of (3.9), the wage is equal to the marginal product of labor:

$$w_t = F_2(K_t, A_t L_t).$$

When the production function has constant returns to scale, it is homogeneous of degree one.[5] Then it follows that production becomes

$$Y_t = K_t F_1(K_t, A_t L_t) + A_t L_t F_2(K_t, A_t L_t).$$

Dividing both sides by Y_t, we obtain that the labor share is one minus the capital share. Therefore, the labor share is also constant when the capital share is constant.

We can also confirm the stability of the labor share by direct calculation. As for the case of r_t, it can be shown, by differentiating $f(K_t/A_t L_t) = F(K_t, A_t L_t)/(A_t L_t)$ with respect to $A_t L_t$, that

$$w_t = f(\tilde{k}_t) - \tilde{k}_t f'(\tilde{k}_t).$$

It can readily be seen that w_t is constant when $\tilde{k}_t$ is constant at $\bar{\tilde{k}}$. Because $A_t L_t / Y_t = 1/f(\tilde{k}_t)$, it is also constant along the balanced growth path. Therefore, $w_t A_t L_t / Y_t$ is also constant. Note that, for a Cobb-Douglas production function $Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}$, the capital share is α and the labor share is $1-\alpha$ regardless of the values of K_t and $A_t L_t$, and therefore the factor shares are constant even outside the balanced-growth path.

[5] Recall that a function $f(x)$ is homogeneous of degree r (is $H(r)$) when $f(sx) = s^r f(x)$ for all (s, x); here x is a vector and r and s are scalars. If $r = 1$, it then follows, using differentiation with respect to s and each element of x, that $f(x) = \sum_i (\partial f / \partial x_i) x_i$ for all x.

3.4 Convergence

We have already seen that, in the Solow model, the economy monotonically converges to the steady state (or balanced growth path). Here, we look at this convergence property more in detail and take a quick look at the data.

3.4.1 Local properties: the speed of convergence

In the basic model, where we again set $L_t = 1$ and use the per-capita notation k_t, the fundamental equation (3.5) can be approximated around the steady-state as

$$\Delta k_{t+1} = \left(1 - \delta + sf'(\bar{k})\right) \Delta k_t, \tag{3.10}$$

where Δk_t represents the deviation of k_t from its steady-state value, that is, $k_t - \bar{k}$, when the deviation is small.[6] When the production function is in the Cobb-Douglas form, using the steady-state solution (3.6),

$$\Delta k_{t+1} = (1 - \delta(1-\alpha))\Delta k_t$$

holds. Replacing Δk_t by $k_t - \bar{k}$ and dividing both sides by $\bar{k}$, (3.10) can be expressed as

$$\frac{k_{t+1} - \bar{k}}{\bar{k}} = (1-\lambda)\frac{k_t - \bar{k}}{\bar{k}},$$

where $\lambda \equiv \delta - sf'(\bar{k})$ represents the *convergence speed*. A large value of λ implies that Δk_{t+1} becomes smaller (in absolute value) more quickly, implying a faster convergence. This is illustrated in Figure 3.5 below: a higher λ represents a flatter slope at the steady state and "more steps until you reach steady state."

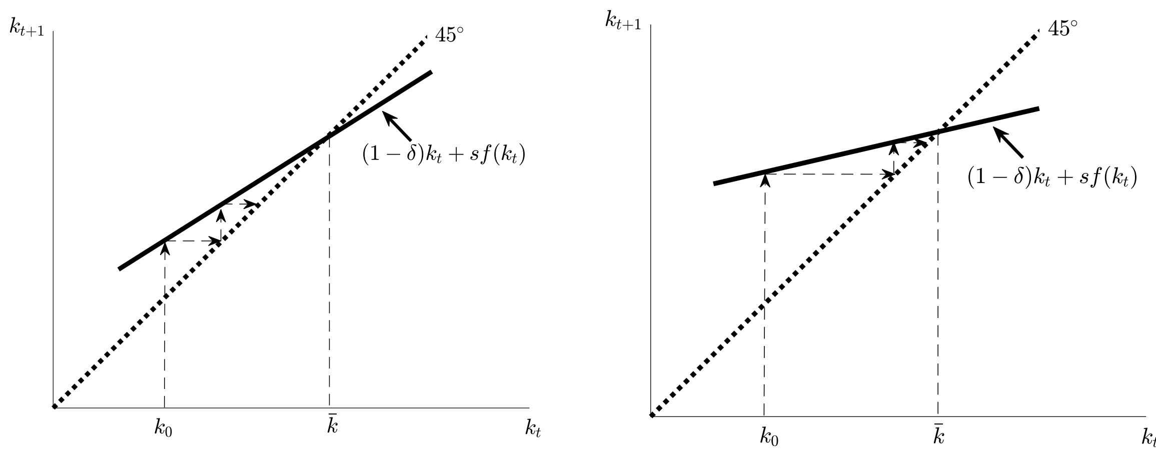

Figure 3.5: Slow and fast convergence.

In the Cobb-Douglas case,

$$\lambda = \delta(1-\alpha) \tag{3.11}$$

[6] This is obtained from a first-order Taylor approximation of the expression around $\bar{k}$.

holds, and the convergence is faster when α is small and δ is large. Note that s does not affect the convergence speed in this case. The convergence speed, in general, is affected by how k_t affects k_{t+1}. In an extreme case, if k_t has no effect on k_{t+1} (a flat line), convergence is immediate. The parameter s has two opposing forces to this mechanism. For a given k_t, a large s implies a larger impact of k_t on k_{t+1}. However, the steady-state value of capital is larger when s is larger, and the marginal product of capital at the steady state, $f'(\bar{k})$, is smaller when s is larger, implying a smaller impact of k_t on k_{t+1}. These two opposing forces exactly offset each other when the production function is in the Cobb-Douglas form.[7]

In the case with growth, the fundamental equation (3.7) can be approximated by

$$(1+\gamma)(1+n)\Delta\tilde{k}_{t+1} = \left(1-\delta+sf'(\bar{\tilde{k}})\right)\Delta\tilde{k}_t. \tag{3.12}$$

When the production function is in the Cobb-Douglas form, using the steady-state solution (3.8),

$$\Delta\tilde{k}_{t+1} = \left(\alpha + \frac{(1-\alpha)(1-\delta)}{(1+\gamma)(1+n)}\right)\Delta\tilde{k}_t$$

holds. The equation (3.12) can be rewritten as

$$\frac{\tilde{k}_{t+1}-\bar{\tilde{k}}}{\bar{\tilde{k}}} = (1-\lambda)\frac{\tilde{k}_t-\bar{\tilde{k}}}{\bar{\tilde{k}}},$$

where the convergence speed is now given by $\lambda = 1-(1-\delta+sf'(\bar{\tilde{k}}))/(1+\gamma)(1+n)$. In the Cobb-Douglas case we obtain

$$\lambda = (1-\alpha)\left(1-\frac{1-\delta}{(1+\gamma)(1+n)}\right). \tag{3.13}$$

3.4.2 Cross-country data

Is convergence observed in the data? Recall that the convergence prediction implies that a country that starts with a smaller per-capita GDP experiences faster subsequent growth. Figure 3.6 plots this relationship across countries. The data is taken from the Penn World Table 10.0 (https://www.rug.nl/ggdc/productivity/pwt/). The horizontal axis is per-capita real GDP in 1960, which we take as the starting point. The vertical axis is the subsequent growth rate (annualized using geometric averages) in per-capita real GDP from 1960 to 2019.

One can immediately see that there is no systematic tendency for initially poor countries to grow faster. The fact that there is no tendency for countries to converge, however, does not imply a rejection of the Solow model. In fact, the Solow model does not predict that different countries will always converge to the same balanced growth path (a phenomenon called "unconditional convergence" or "absolute convergence"). Rather, it predicts that countries converge if they share the same parameter values ("conditional convergence"). We

[7] The Cobb-Douglas function is very convenient because it often simplifies the algebra and leads to simple expressions. This simplicity, however, can be deceiving as we see here: the functional form often makes fundamental forces going in opposite direction cancel. That is, under the surface there may be very strong forces but, as if by magic, the Cobb-Douglas form makes them invisible.

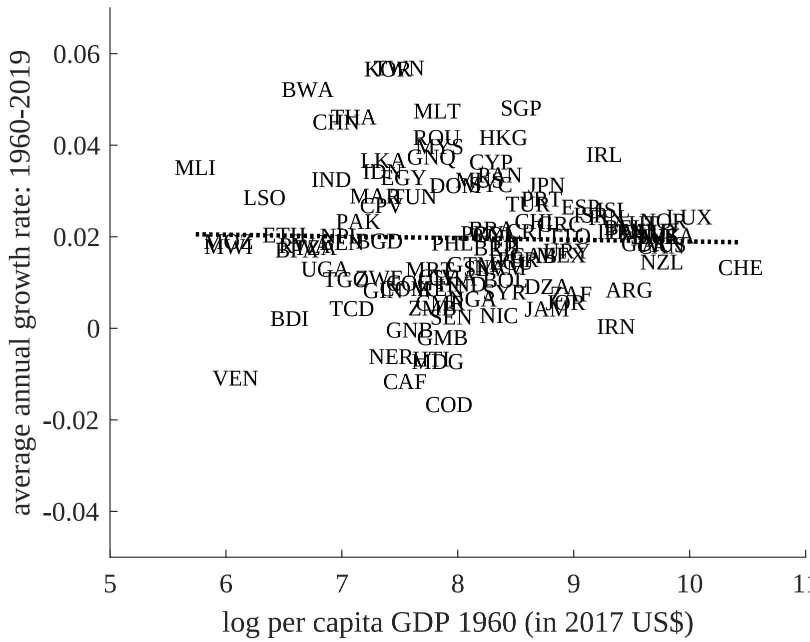

Figure 3.6: All countries, 1960–2019.

Source: Penn World Tables 10.0. The GDP variable used is RGDPNA.

know, for example, that saving rates differ widely across countries and, at least over shorter time horizons, it is reasonable to think that the growth rates of A_t also differ.

To examine conditional convergence, a useful exercise is to look at the same kind of graph restricted to a smaller, and more similar, set of countries. Figure 3.7 plots the same data as Figure 3.6, but only for the original members of the Organisation for Economic Co-operation and Development (OECD). OECD was formed by high-income countries that share relatively similar economic and political institutions, and we can expect the underlying parameters in the Solow model to be relatively similar among these countries.

For this set of countries, we observe a clear tendency for convergence: poor countries in 1960 on average experience faster subsequent growth. Barro and Sala-i-Martin (1995) (Chapter 12) conduct a similar exercise across regions within countries, treating each region as a different "country." For U.S. states, Japanese prefectures, and European regions, they find a clear tendency of convergence, supporting the prediction of the Solow model.

In a recent paper, Kremer, Willis, and You (2022) show that in recent years, the data actually show a tendency for unconditional convergence. Figure 3.8 below repeats the same exercise as in Figure 3.6 for the same set of countries, but setting the initial date to 2000. We can see that some convergence (negative correlation) is observed in the recent years. The authors argue that this tendency arose because some of the underlying factors that affect growth (the factors that likely affect the growth rate of A_t), such as policies, institutions, and human capital have become more similar across countries in recent years. In Chapter 13, we will discuss this and many related issues in greater detail.

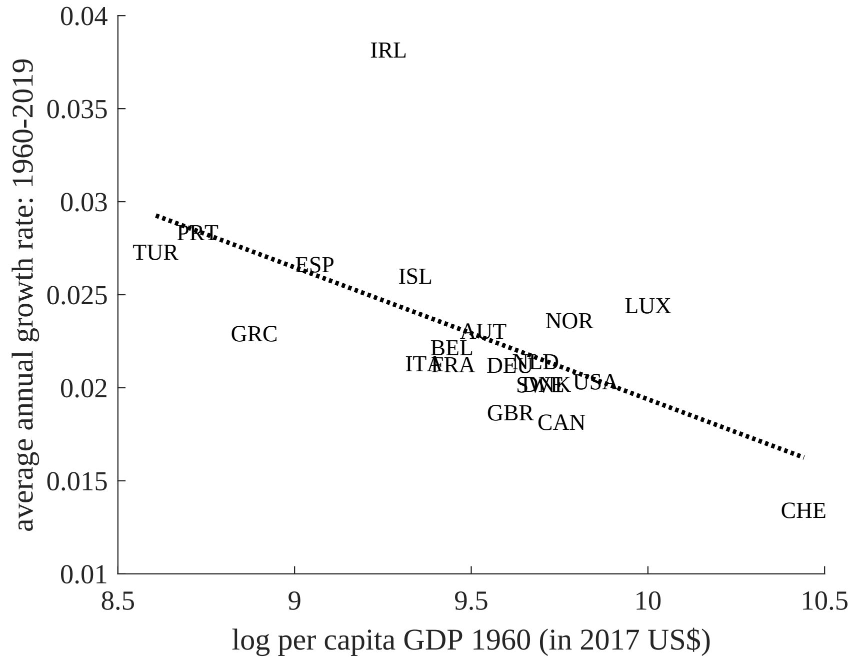

Figure 3.7: OECD countries, 1960–2019.

Source: Penn World Tables 10.0. The GDP variable used is RGDPNA.

3.4.3 Quantitative use of the Solow model

What are the *quantitative* predictions of the model for convergence? To answer this question, we need to assign functional forms and specific parameter values. This procedure will give us a numerical value for λ in (3.13). Once λ is computed, one can of course also conduct counterfactual experiments by simulating a hypothetical situation using the quantitative model.

If we are interested in the model's quantitative predictions for the speed of convergence, one way to proceed would be to simply see if it is possible to choose parameters so as to hit the "observed λ" (e.g., as measured by the slope in Figure 3.7). More generally, one could specify a full stochastic model, say, with explicit shocks to variables (such as A_t) and estimate the resulting structure against the data we just looked at. Clearly, we could generate a good fit in this case if we are free to choose parameters; for example, given any production function, we could match the λ by an appropriate choice of δ. This choice, however, may not be consistent with what we know about depreciation rates from microeconomic data. More generally, we would like our model's different components (functional forms and parameter values) to be selected to be in line with microeconomic studies (and perhaps aggregate data too). This way, the quantitative evaluation is disciplined. As briefly discussed in Chapter 1 above, a way forward is *calibration*.[8] The procedure consists of two distinct steps, each guided by data. First, we assign a parameterized functional forms to the unknown functions

[8]We describe calibration, and compare it to other methods, in much more detail in Chapter 8.

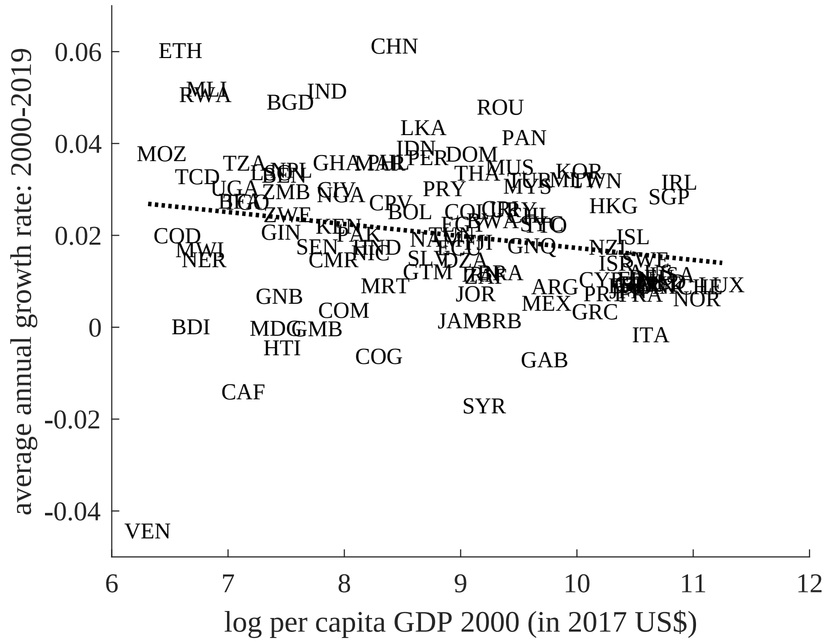

Figure 3.8: All countries, 2000–2019.

Source: Penn World Tables 10.0. The GDP variable used is RGDPNA.

in the model. In our case, the production function F corresponds to the unknown function. Second, we assign specific values to the parameters. Because we use various moments (such as means and variances) in the data to assign parameter values, this second step bears close resemblance to estimation by the method of moments.

Before starting the calibration, we have to decide on the length of the time period. Here, we are chiefly interested in movements in aggregates that occur over the medium run and set one period to be one year. For the first step, we choose the Cobb-Douglas form, used above, for the production function: $F(K_t, A_tL_t) = K_t^{\alpha}(A_tL_t)^{1-\alpha}$. As we have seen, this functional form yields constant factor shares, which is consistent with the rather striking data on an absence of major movements in the shares. A Cobb-Douglas function is special, however, in that it has the property that the substitution elasticity between inputs (capital and labor) is equal to one, so we need to make sure that it is consistent with studies of production functions, at least at a high level of aggregation.[9] These studies rarely suggest major departures from 1.

In the second step, we need to assign values to the parameters α, δ, γ, n, and s. Here, we have implicitly assumed that the production function has constant returns to scale, i.e., that the sum of the exponents on capital and labor is one, so that it suffices to choose the α. Before discussing the parameter choices, note that for the convergence speed λ in (3.13), the information on s is not necessary. Below we will therefore skip assigning a value to s.

[9]The substitution elasticity is given by the percentage change in the ratio of the inputs when the ratio of their prices change by one percent, i.e., $-d\log(K/L)/d\log(r/w)$. Using the firm's first-order conditions, we can see that this expression must equal 1 for a Cobb-Douglas production function.

Calibration usually draws on multiple data sources. Overall, there are two methods of assigning the parameter values based on data. First, if particular parameters have been considered as important objects for investigation elsewhere and we know plausible parameter values from past studies, it is convenient to directly assign the parameter values accordingly. Second, we can choose data moments that involve several parameters and assign parameter values such that these moments, when generated by our theory, line up with observed moments. The first method can be considered a special case of the second method, because the parameter values from past studies have to come from certain data moments used in these studies. From this perspective, an alternative interpretation of the calibration procedure is an implementation of the method of moments with multiple data sets.

We set the value of α from national income accounting. As we have seen from the previous section, α corresponds to the capital share. From Figure 2.12 in Chapter 2, $\alpha = 1/3$ is a good approximation. We can set the value of γ from the long-run growth rate of per capita income in advanced countries. The population growth rate n can be measured directly from the data. Barro and Sala-i-Martin (2004) use $\gamma = 0.02$ and $n = 0.01$ as a benchmark at the annual frequency. When $\tilde{k}_{t+1} = \tilde{k}_t$ (along the balanced growth path), the fundamental equation (3.7) can be rewritten as

$$(1+\gamma)(1+n) = 1 - \delta + \frac{I_t}{K_t},$$

where we used $sf(\tilde{k}_t)/\tilde{k}_t = I_t/K_t$. The investment-capital ratio in the U.S. economy is about 0.076 (Cooley and Prescott, 1995) at an annual frequency. Given the above values of γ and n, this equation implies $\delta = 0.046$.[10] Clearly, here, one could alternatively have used depreciation rates directly from data on depreciation (by capital type, or from aggregate data on depreciation rates) and then this equation would have implied a value for the average size of I/K on a balanced growth path. Given the accounting practices and the fact that capital remains at roughly three times annual GDP as time passes, on average investment precisely will have to make up for depreciation (taking population and technology growth into account), so either way the number for δ ends up around 0.05.

With these values, we find that $\lambda = 0.049$. The empirical counterpart of λ is about 0.015 to 0.03 (Barro and Sala-i-Martin 2004, p,59), and therefore the Solow model over-predicts the convergence speed. The model value of λ is about 0.02 if α is raised to 0.73. An argument for a larger value of α is that a part of labor income is the return to human capital, which can be accumulated in a similar manner as physical capital, and some part of labor income should be included in the capital share. Relatedly, one can view A_t as an accumulable factor—after all, technological development is often part of conscious investments into R&D, and hence it too is a capital stock. These issues are returned to in Chapter 13.

Once the model has been assigned specific functional forms and parameter values, we can also conduct quantitative experiments. Suppose, for example, the saving rate s equals 0.1. How would increasing s to 0.2 affect the normalized level of output along the balanced-growth path, $f(\bar{\tilde{k}})$? With a Cobb-Douglas production function, we have already obtained

[10] Given that $I/K = (I/Y)(Y/K)$, we could alternatively have measured I/Y—the saving rate—and Y/K, which we know is about 1/3. Conversely, we can now obtain the implied I/Y as 0.076/0.33, which approximately equals 0.23.

the solution for $\bar{\tilde{k}}$ in (3.8). Inserting our calibrated parameter values, along with $s = 0.1$, we obtain $\bar{\tilde{k}} = 1.20$ and $f(\bar{\tilde{k}}) = \bar{\tilde{k}}^{\alpha} = 1.06$. When s goes up to 0.2, $\bar{\tilde{k}}$ rises to 1.90 and $f(\bar{\tilde{k}}) = 1.24$. Therefore, doubling the saving rate from 10% to 20% increases the normalized level of output by 17%, because $1.24/1.06 = 1.17$. In addition, we could compute the quantitative predictions for how fast output would rise to eventually reach a 17% higher value.

3.5 Business cycles

The usefulness of the Solow model goes much beyond the study of economic growth; due to its ability to account for the broad features of the macroeconomic data over our modern economic history, it constitutes the core of macroeconomic modeling. A prominent illustration of this is the fact that virtually all modern theories of business cycles build on a version of, or elaboration on, the Solow model. Although business cycle theories are detailed later in Chapters 14 and 18, we now briefly review how these theories are linked to the Solow model and exhibit some of the key associated tools, such as impulse response diagrams.

3.5.1 Various theories of business cycles

The studies of business cycles are primarily the analysis of the arrival of shocks to the economy and how the economy reacts to these shocks. The way the economy reacts to the shocks is usually called the "propagation mechanism." Once we introduce uncertainty in Chapter 7, we can treat these "shocks" more precisely. Here, we consider a general (exogenous and deterministic) movement in certain variables as the source of business cycle fluctuations.

Below, first, we extend the basic model in several directions. In particular, we outline how the Solow model can be modified and accommodate various shocks in three different business cycle models. Throughout we conduct the analysis in per-capita terms and, hence, use lower-case letters.

- The first business cycle model is the so-called real business cycle (RBC) model. As will be explained in Chapter 14, the RBC model provides a simple mechanism whereby macroeconomic variables comove, as is clear in the data. The prototypical RBC model considers the following aggregate production function

  $$y_t = A_t F(k_t, \ell_t),$$

  where ℓ_t is variable labor input, and considers a shock to A_t (often called the "neutral technology shock"). That is, it assumes that A_t changes over time and the movement of A_t is the source of the business cycle. For example, consider a model where A_t switches around between two values, A_H and A_L, where $A_H > A_L$. When A_t moves to A_H, the economy starts moving towards the corresponding steady state. Then A_t switches to A_L, and the economy now moves towards a different steady-state value. We can interpret these movements as business cycles: movements around some average. Augmented with steady growth we would have movements around the balanced path.

Note that the production function in this section is different from the Harrod-neutral form $F(k_t, A_t\ell_t)$ earlier. Note also that the distinction between Harrod-neutral technological progress and the Hicks-neutral technological progress is not essential when the production function is in the Cobb-Douglas form: the Harrod-neutral production function $k_t^\alpha(A_t\ell_t)^{1-\alpha}$ can be rewritten as $A_t^{1-\alpha}k_t^\alpha\ell_t^{1-\alpha}$, and by defining $\tilde{A}_t \equiv A_t^{1-\alpha}$, the same production can be interpreted as the Hicks-neutral production function $\tilde{A}_t k_t^\alpha \ell_t^{1-\alpha}$.

If we maintain the assumptions of the basic model, we have ℓ_t constant and $i_t/y_t = s$ (and $c_t/y_t = 1 - s$) constant even with the shocks to A_t. These features are at odds with the business cycle data. To accommodate the business cycle facts that (i) ℓ_t comoves positively with the business cycle, (ii) i_t is more volatile than y_t, and (iii) c_t is less volatile than y_t, the basic equation would have to allow for the saving rate and ℓ_t to react to A_t (and possibly to k_t). The economy evolves, therefore, following the difference equation

$$k_{t+1} = (1-\delta)k_t + s(k_t, A_t)A_tF(k_t, \ell(k_t, A_t)).$$

This is a modified form of the fundamental equation (3.5) of the basic Solow model.

- Next, we consider a different kind of shock. Suppose that the final goods market clearing condition (3.3) is modified to

$$y_t = c_t + i_t/\nu_t,$$

where ν_t moves over time (and often called the "investment-specific technological progress"): when it is high, it is cheaper to produce investment goods. One can think of this equation as reflecting the two-sector structure of the economy: y_t and c_t are measured in consumption goods, and investment goods have a production process that can create i_t units of investment goods by using i_t/ν_t units of consumption goods. The fundamental equation (3.5) can now be modified to

$$k_{t+1} = s\nu_tF(k_t, \ell) + (1-\delta)k_t;$$

this equation can also be extended to include endogenous s and ℓ as in the case of the neutral technology shock.

- In the third example, we consider a very different model structure: one with "demand shocks." First, assume that c_t is exogenous and that it moves around over time. The movement of c_t serves as the (demand) shock. Suppose, further, that we maintain the assumption that $i_t/c_t = s/(1-s)$. Because of this assumption, since s is constant, i_t is proportional to c_t and moves along with it. Therefore, the total demand for goods is a function of c_t:

$$c_t + i_t = \frac{1}{1-s}c_t.$$

When c_t is not sufficiently large, $c_t + i_t$ would be less than the full capacity output $F(k_t, \ell)$ (here we assume again that ℓ is given by labor-force participation: those who want to work). Assume, then, that when there are demand shortages, total output y_t

is determined by the demand side and so that a fraction u_t of the labor force become unemployed (therefore, u_t is the unemployment rate):

$$y_t = \frac{1}{1-s} c_t = F(k_t, \ell(1-u_t)).$$

From the second equality, u_t can be represented as the function of c_t and k_t: $u(c_t, k_t)$. The fundamental equation (3.5) can therefore be modified as

$$k_{t+1} = sF(k_t, \ell(1-u(c_t, k_t))) + (1-\delta)k_t.$$

This framework is very Keynesian in spirit but clearly begs the question of how output can end up below full capacity, and hence be driven by demand. In a well-functioning market, this phenomenon could not occur. This book contains several chapters on frictions that could lead to something like the setting just described, and it then becomes central for policymakers to understand the exact nature of these frictions.[11]

In all three cases above, we can represent k_{t+1} as a function of k_t and a shock (A_t, ν_t, or c_t). This representation allows us to characterize the dynamics of k_t (and other macroeconomic variables, such as y_t, c_t, and i_t) in response to these shocks.

3.5.2 Impulse responses

One method of describing how the economic variables respond to shocks is to draw an impulse-response function. Consider the RBC example above, with a Cobb-Douglas production function, an exogenous saving rate, and fixed labor supply. Suppose that before period 0, the value of A_t is constant at $\bar{A}$. After a sufficiently long time, the value of k_t settles close to the corresponding steady-state value $\bar{k}$. Then suppose that at time 0, A_0 is $(\varepsilon \times 100)\%$ higher, that is, $A_0 = (1+\varepsilon)\bar{A}$. For $t = 1, 2, 3, \ldots$, the value of A_t is $A_t = (1 + \rho^t \varepsilon)\bar{A}$, where $\rho \in (0, 1)$.

We can then generate the resulting time-path of k_t, starting from $k_0 = \bar{k}$, with

$$k_{t+1} = sA_t k_t^{\alpha} \ell^{1-\alpha} + (1-\delta)k_t, \tag{3.14}$$

for $t = 0, 1, 2, \ldots$. This time path is called the impulse-response function. The time-path of A_t is drawn in Figure 3.9—the impulse—along with the response of k_t. For the impulse-response function for k_t we use $s = 0.2$, $\delta = 0.046$, and $\alpha = 1/3$. The starting value of A and the value of ℓ are normalized to 1. The initial value of the shock to A, ε, is 1%, and the persistence $\rho = 0.9$.

In the figure, k_t increases from its steady-state value of 9.066 to a peak of 9.089, then gradually returns to its original steady-level—producing a hump-shaped response. This occurs because capital adjusts gradually in response to shocks: the impact of a temporary rise in A_t raises output and investment, which in turn increases capital accumulation, but only with a delay. Three features of the graph are worth emphasizing. First, the adjustment

[11] Chapter 21, in particular, describes a related example where c is a "production externality" and, as such, acts like a demand channel.

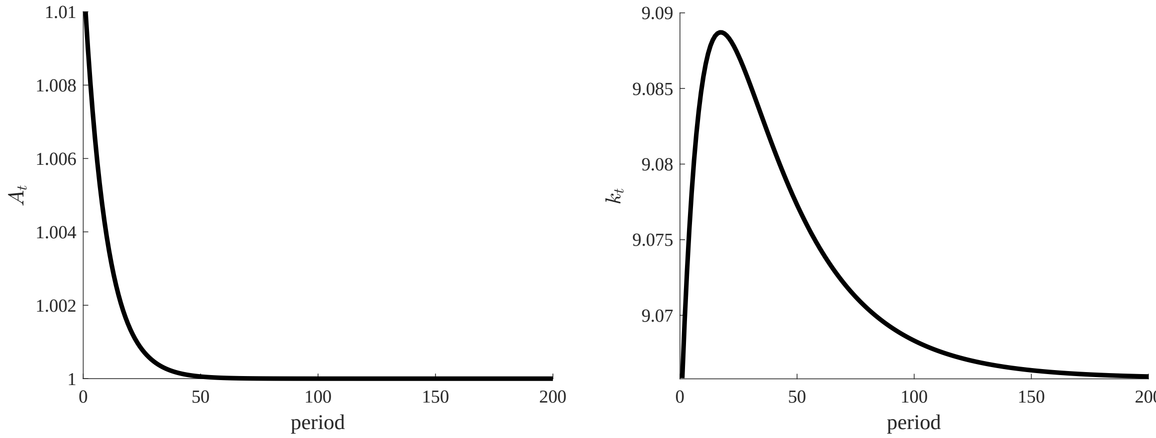

Figure 3.9: Impulse response: how A_t (left) affects k_t (right).

of k_t is much slower than the movement of A_t; while the deviation of A_t fades rapidly and becomes negligible around period 50, the response of k_t is far more persistent. Second, the magnitude of the capital response is small relative to the impulse: the maximum deviation, measured as a fraction to the steady-state level is $9.089/9.066 - 1 = 0.0025$, or just 0.25%, compared to the initial 1% deviation in A_t. Third, k_t eventually returns to its steady-state value. This reflects the convergence force inherent in the model, which remains active even in the presence of recurrent shocks, as in the business-cycle examples.

Log-linearized impulse responses

Often it is more convenient to approximate the dynamics around the steady state by a log-linarized system. Log-linearization expresses the system in terms of deviation of the logarithms of variables from their corresponding steady-state values. An arbitrary variable x_t can be expressed as

$$x_t = \bar{x}e^{\hat{x}_t}, \tag{3.15}$$

where

$$\hat{x}_t \equiv \log\left(\frac{x_t}{\bar{x}}\right),$$

is the log deviation from the steady-state value $\bar{x}$. Note that because

$$\hat{x}_t \equiv \log\left(\frac{x_t}{\bar{x}}\right) \approx \frac{x_t - \bar{x}}{\bar{x}},$$

where the approximation is the first-order Taylor expansion around $\bar{x}$, $\hat{x}_t$ can be interpreted as the percent deviation from the steady state. Also note that

$$\bar{x}e^{\hat{x}_t} \approx \bar{x}(1 + \hat{x}_t) \tag{3.16}$$

from the first-order Taylor approximation of the expression.

Consider the impulse-response experiment of the system (3.14). Using the transformation

(3.15),

$$\bar{k}e^{\hat{k}_{t+1}} = s\bar{A}e^{\hat{A}_t}(\bar{k}e^{\hat{k}_t})^\alpha \ell^{1-\alpha} + (1-\delta)\bar{k}e^{\hat{k}_t}$$

holds. Using (3.16) and the fact that

$$\bar{k} = s\bar{A}(\bar{k})^\alpha \ell^{1-\alpha} + (1-\delta)\bar{k} \tag{3.17}$$

holds from the definition of $\bar{k}$, we obtain

$$\bar{k}\hat{k}_{t+1} = s\bar{A}(\bar{k})^\alpha \ell^{1-\alpha}(\hat{A}_t + \alpha\hat{k}_t) + (1-\delta)\bar{k}\hat{k}_t.$$

Because equation (3.17) implies $\delta\bar{k} = s\bar{A}(\bar{k})^\alpha \ell^{1-\alpha}$, this equation can be rewritten as

$$\hat{k}_{t+1} = (1-\delta(1-\alpha))\hat{k}_t + \delta\hat{A}_t. \tag{3.18}$$

Let us again use $\lambda \equiv \delta(1-\alpha)$, which is the notation for the convergence speed in (3.11). In fact, the log-linearization procedure here is essentially the same as the procedure for obtaining the percentage deviation in the convergence section, in the sense that both are applying the first-order Taylor approximation to (3.14).

The above impulse-response experiment corresponds to setting $\hat{A}_0 = \varepsilon$ and $\hat{A}_t = \rho^t\varepsilon$ for $t = 1, 2, 3, \ldots$. Solving (3.18) with this specification yields

$$\hat{k}_{t+1} = \sum_{\tau=0}^{t} \rho^{t-\tau}(1-\lambda)^\tau \delta\varepsilon = \rho^t \frac{1 - \left(\frac{1-\lambda}{\rho}\right)^{t+1}}{1 - \frac{1-\lambda}{\rho}}\delta\varepsilon.$$

Quantitatively, the log-linear approximation performs well in our calibrated model. The maximum error (in the units of K_t) of the approximation is about 0.00001. If we were to plot the log-linear solution and the nonlinear solution in the same figure, the difference would not be visible.

Applying the above log-linearization procedure to the production function $y_t = A_t k_t^\alpha \ell^{1-\alpha}$, the log-deviation of output is

$$\hat{y}_t = \hat{A}_t + \alpha\hat{k}_t = \rho^t\delta\varepsilon + \alpha\rho^{t-1}\frac{1 - \left(\frac{1-\lambda}{\rho}\right)^t}{1 - \frac{1-\lambda}{\rho}}\delta\varepsilon = \rho^t\left(1 + \frac{\alpha}{\rho}\frac{1 - \left(\frac{1-\lambda}{\rho}\right)^t}{1 - \frac{1-\lambda}{\rho}}\right)\delta\varepsilon.$$

for $t = 1, 2, \ldots$ (for $t = 0$, $\hat{y}_0 = \hat{A}_0 = \varepsilon$). This equation explicitly describes how y_t moves over the cycle, in response to the realization of the shock ε. Log-linearization allows us to derive this explicit expression.

An alternative to the log-linear approximation is a linear approximation in *levels*. Let the production function be $A_t F(k_t, \ell_t)$. Assuming that $\ell_t = 1$ for any t and defining $f(k_t) \equiv F(k_t, 1)$, the fundamental equation (3.5) (with the modification in the production function) can be written as

$$k_{t+1} = g(k_t, A_t),$$

where $g(k_t, A_t) \equiv (1-\delta)k_t + sA_t f(k_t)$. Using the notation we employed earlier,

$$\Delta k_{t+1} = g_k \Delta k_t + g_A \Delta A_t,$$

where g_k and g_A are the partial derivatives of $g(k, A)$ with respect to k and A, respectively (evaluated at the steady state), and $\Delta k_t = k_t - \bar{k}$ and $\Delta A_t = A_t - \bar{A}$. When $f(k_t) = k_t^\alpha$, we obtain $g_k = 1 - \delta(1-\alpha)$ and $g_A = s^{\frac{1}{1-\alpha}}\delta^{-\frac{\alpha}{1-\alpha}}A^{\frac{\alpha}{1-\alpha}}$.