2

The Solow Growth Model

The previous chapter introduced a number of basic facts and posed the main questions concerning the sources of economic growth over time and the causes of differences in economic performance across countries. These questions are central not only to growth theory but also to macroeconomics and the social sciences more generally. Our next task is to develop a simple framework that can help us think about the proximate causes and the mechanics of the process of economic growth and cross-country income differences. We will use this framework both to study potential sources of economic growth and also to perform simple comparative statics to gain an understanding of which country characteristics are conducive to higher levels of income per capita and more rapid economic growth.

Our starting point is the so-called Solow-Swan model named after Robert (Bob) Solow and Trevor Swan, or simply the Solow model, named after the more famous of the two economists. These economists published two pathbreaking articles in the same year, 1956 (Solow, 1956; Swan, 1956) introducing the Solow model. Bob Solow later developed many implications and applications of this model and was awarded the Nobel prize in economics for his contributions. This model has shaped the way we approach not only economic growth but also the entire field of macroeconomics. Consequently, a by-product of our analysis of this chapter is a detailed exposition of a workhorse model of macroeconomics.

The Solow model is remarkable in its simplicity. Looking at it today, one may fail to appreciate how much of an intellectual breakthrough it was. Before the advent of the Solow growth model, the most common approach to economic growth built on the model developed by Roy Harrod and Evsey Domar (Harrod, 1939; Domar, 1946). The Harrod-Domar model emphasized potential dysfunctional aspects of economic growth, for example, how economic growth could go hand-in-hand with increasing unemployment (see Exercise 2.23 on this model). The Solow model demonstrated why the Harrod-Domar model was not an attractive place to start. At the center of the Solow growth model, distinguishing it from the Harrod-Domar model, is the neoclassical aggregate production function. This function not only enables the Solow model to make contact with microeconomics, but as we will see in the next chapter, it also serves as a bridge between the model and the data.

An important feature of the Solow model, which is shared by many models presented in this book, is that it is a simple and abstract representation of a complex economy. At first, it may appear too simple or too abstract. After all, to do justice to the process of growth or macroeconomic equilibrium, we have to consider households and individuals with different tastes, abilities, incomes, and roles in society; various sectors; and multiple social interactions. The Solow model cuts through these complications by constructing a simple one-good economy, with little reference to individual decisions. Therefore, the Solow model should be thought of as a starting point and a springboard for richer models.

In this chapter, I present the basic Solow model. The closely related neoclassical growth model is presented in Chapter 8.

2.1 The Economic Environment of the Basic Solow Model

Economic growth and development are dynamic processes and thus necessitate dynamic models. Despite its simplicity, the Solow growth model is a dynamic general equilibrium model (though, importantly, many key features of dynamic general equilibrium models emphasized in Chapter 5, such as preferences and dynamic optimization, are missing in this model).

The Solow model can be formulated in either discrete or continuous time. I start with the discrete-time version, because it is conceptually simpler and more commonly used in macroeconomic applications. However, many growth models are formulated in continuous time, and I then provide a detailed exposition of the continuous-time version of the Solow model and show that it is often more convenient to work with.

2.1.1 Households and Production

Consider a closed economy, with a unique final good. The economy is in discrete time running to an infinite horizon, so that time is indexed by $t = 0, 1, 2, \ldots$. Time periods here may correspond to days, weeks, or years. For now, we do not need to specify the time scale.

The economy is inhabited by a large number of households. Throughout the book I use the terms *households, individuals,* and *agents* interchangeably. The Solow model makes relatively few assumptions about households, because their optimization problem is not explicitly modeled. This lack of optimization on the household side is the main difference between the Solow and the *neoclassical growth* models. The latter is the Solow model plus dynamic consumer (household) optimization. To fix ideas, you may want to assume that all households are identical, so that the economy trivially admits *a representative household*—meaning that the demand and labor supply side of the economy can be represented as if it resulted from the behavior of a single household. The representative household assumption is discussed in detail in Chapter 5.

What do we need to know about households in this economy? The answer is: not much. We have not yet endowed households with preferences (utility functions). Instead, for now, households are assumed to save a constant exogenous fraction $s \in (0, 1)$ of their disposable income—regardless of what else is happening in the economy. This assumption is the same as that used in basic Keynesian models and the Harrod-Domar model mentioned above. It is also at odds with reality. Individuals do not save a constant fraction of their incomes; if they did, then an announcement by the government that there will be a large tax increase next year should have no effect on their savings decisions, which seems both unreasonable and empirically incorrect. Nevertheless, the exogenous constant saving rate is a convenient starting point, and we will spend a lot of time in the rest of the book analyzing how consumers behave and make intertemporal choices.

The other key agents in the economy are firms. Firms, like consumers, are highly heterogeneous in practice. Even within a narrowly defined sector of an economy, no two firms are identical. But again for simplicity, let us start with an assumption similar to the representative household assumption, but now applied to firms: suppose that all firms in this economy have access to the same production function for the final good, or that the economy admits a *representative firm,* with a representative (or aggregate) production function. The conditions under which this represetive firm assumption is reasonable are also discussed in Chapter 5. The aggregate production function for the unique final good is written as

$$Y(t) = F(K(t), L(t), A(t)), \tag{2.1}$$

where $Y(t)$ is the total amount of production of the final good at time t, $K(t)$ is the capital stock, $L(t)$ is total employment, and $A(t)$ is technology at time t. Employment can be measured in different ways. For example, we may want to think of $L(t)$ as corresponding to hours of employment or to number of employees. The capital stock $K(t)$ corresponds to the quantity of "machines" (or more specifically, equipment and structures) used in production, and it is typically measured in terms of the value of the machines. There are also multiple ways of thinking of capital (and equally many ways of specifying how capital comes into existence). Since the objective here is to start with a simple workable model, I make the rather sharp simplifying assumption that capital is the same as the final good of the economy. However, instead of being consumed, capital is used in the production process of more goods. To take a concrete example, think of the final good as "corn." Corn can be used both for consumption and as an input, as seed, for the production of more corn tomorrow. Capital then corresponds to the amount of corn used as seed for further production.

Technology, on the other hand, has no natural unit, and $A(t)$ is simply a *shifter* of the production function (2.1). For mathematical convenience, I often represent $A(t)$ in terms of a number, but it is useful to bear in mind that, at the end of the day, it is a representation of a more abstract concept. As noted in Chapter 1, we may often want to think of a broad notion of technology, incorporating the effects of the organization of production and of markets on the efficiency with which the factors of production are utilized. In the current model, $A(t)$ represents all these effects.

A major assumption of the Solow growth model (and of the neoclassical growth model we will study in Chapter 8) is that technology is *free:* it is publicly available as a nonexcludable, nonrival good. Recall that a good is *nonrival* if its consumption or use by others does not preclude an individual's consumption or use. It is *nonexcludable*, if it is impossible to prevent another person from using or consuming it. Technology is a good candidate for a nonexcludable, nonrival good; once the society has some knowledge useful for increasing the efficiency of production, this knowledge can be used by any firm without impinging on the use of it by others. Moreover, it is typically difficult to prevent firms from using this knowledge (at least once it is in the public domain and is not protected by patents). For example, once the society knows how to make wheels, everybody can use that knowledge to make wheels without diminishing the ability of others to do the same (thus making the knowledge to produce wheels nonrival). Moreover, unless somebody has a well-enforced patent on wheels, anybody can decide to produce wheels (thus making the knowhow to produce wheels nonexcludable). The implication of the assumptions that technology is nonrival and nonexcludable is that $A(t)$ is freely available to all potential firms in the economy and firms do not have to pay for making use of this technology. Departing from models in which technology is freely available is a major step toward understanding technological progress and will be our focus in Part IV.

As an aside, note that some authors use x_t or K_t when working with discrete time and reserve the notation $x(t)$ or $K(t)$ for continuous time. Since I go back and forth between continuous and discrete time, I use the latter notation throughout. When there is no risk of confusion, I drop the time arguments, but whenever there is the slightest risk of confusion, I err on the side of caution and include the time arguments.

Let us next impose the following standard assumptions on the aggregate production function.

Assumption 1 (Continuity, Differentiability, Positive and Diminishing Marginal Products, and Constant Returns to Scale) *The production function* $F : \mathbb{R}^3_+ \to \mathbb{R}_+$ *is twice differentiable in* K *and* L*, and satisfies*

$$F_K(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial K} > 0, \quad F_L(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial L} > 0,$$

$$F_{KK}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial K^2} < 0, \quad F_{LL}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial L^2} < 0.$$

Moreover, F *exhibits constant returns to scale in* K *and* L*.*

All of the components of Assumption 1 are important. First, the notation $F : \mathbb{R}^3_+ \to \mathbb{R}_+$ implies that the production function takes nonnegative arguments (i.e., $K, L \in \mathbb{R}_+$) and maps to nonnegative levels of output ($Y \in \mathbb{R}_+$). It is natural that the level of capital and the level of employment should be positive. Since A has no natural units, it could have been negative. But there is no loss of generality in restricting it to be positive. The second important aspect of Assumption 1 is that F is a continuous function in its arguments and is also differentiable. There are many interesting production functions that are not differentiable, and some interesting ones that are not even continuous. But working with differentiable functions makes it possible to use differential calculus, and the loss of some generality is a small price to pay for this convenience. Assumption 1 also specifies that marginal products are positive (so that the level of production increases with the amount of inputs); this restriction also rules out some potential production functions and can be relaxed without much complication (see Exercise 2.8). More importantly, Assumption 1 requires that the marginal products of both capital and labor are diminishing, that is, $F_{KK} < 0$ and $F_{LL} < 0$, so that more capital, holding everything else constant, increases output by less and less. And the same applies to labor. This property is sometimes also referred to as "diminishing returns" to capital and labor. The degree of diminishing returns to capital plays a very important role in many results of the basic growth model. In fact, the presence of diminishing returns to capital distinguishes the Solow growth model from its antecedent, the Harrod-Domar model (see Exercise 2.23).

The other important assumption is that of constant returns to scale. Recall that F exhibits *constant returns to scale* in K and L if it is *linearly homogeneous* (homogeneous of degree 1) in these two variables. More specifically:

Definition 2.1 *Let* $K \in \mathbb{N}$*. The function* $g : \mathbb{R}^{K+2} \to \mathbb{R}$ *is* homogeneous of degree m *in* $x \in \mathbb{R}$ *and* $y \in \mathbb{R}$ *if*

$$g(\lambda x, \lambda y, z) = \lambda^m g(x, y, z) \text{ for all } \lambda \in \mathbb{R}_+ \text{ and } z \in \mathbb{R}^K.$$

It can be easily verified that linear homogeneity implies that the production function F is concave, though not strictly so (see Exercise 2.2). Linearly homogeneous (constant returns to scale) production functions are particularly useful because of the following theorem.

Theorem 2.1 (Euler's Theorem) *Suppose that* $g : \mathbb{R}^{K+2} \to \mathbb{R}$ *is differentiable in* $x \in \mathbb{R}$ *and* $y \in \mathbb{R}$*, with partial derivatives denoted by* g_x *and* g_y*, and is homogeneous of degree* m *in* x *and* y*. Then*

$$mg(x, y, z) = g_x(x, y, z)x + g_y(x, y, z)y \text{ for all } x \in \mathbb{R}, y \in \mathbb{R}, \text{ and } z \in \mathbb{R}^K.$$

Moreover, $g_x(x, y, z)$ *and* $g_y(x, y, z)$ *are themselves homogeneous of degree* $m - 1$ *in* x *and* y*.*

Proof. We have that g is differentiable and

$$\lambda^m g(x, y, z) = g(\lambda x, \lambda y, z). \tag{2.2}$$

Differentiate both sides of (2.2) with respect to λ, which gives

$$m\lambda^{m-1} g(x, y, z) = g_x(\lambda x, \lambda y, z)x + g_y(\lambda x, \lambda y, z)y$$

for any λ. Setting $\lambda = 1$ yields the first result. To obtain the second result, differentiate both sides of (2.2) with respect to x:

$$\lambda g_x(\lambda x, \lambda y, z) = \lambda^m g_x(x, y, z).$$

Dividing both sides by λ establishes the desired result. ■

2.1.2 Endowments, Market Structure, and Market Clearing

The previous subsection has specified household behavior and the technology of production. The next step is to specify endowments, that is, the amounts of labor and capital that the economy starts with and who owns these endowments. We will then be in a position to investigate the allocation of resources in this economy. Resources (for a given set of households and production technology) can be allocated in many different ways, depending on the *institutional structure* of the society. Chapters 5–8 discuss how a social planner wishing to maximize a weighted average of the utilities of households might allocate resources, while Part VIII focuses on the allocation of resources favoring individuals who are politically powerful. The more familiar benchmark for the allocation of resources is to assume a specific set of market institutions, in particular, competitive markets. In competitive markets, households and firms act in a price-taking manner and pursue their own objectives, and prices clear markets. Competitive markets are a natural benchmark, and I start by assuming that all goods and factor markets are competitive. This is yet another assumption that is not totally innocuous. For example, both labor and capital markets have imperfections, with certain important implications for economic growth, and monopoly power in product markets plays a major role in Part IV. But these implications can be best appreciated by starting out with the competitive benchmark.

Before investigating trading in competitive markets, let us also specify the ownership of the endowments. Since competitive markets make sense only in the context of an economy with (at least partial) private ownership of assets and the means of production, it is natural to suppose that factors of production are owned by households. In particular, let us suppose that households own all labor, which they supply inelastically. Inelastic supply means that there is some endowment of labor in the economy, for example, equal to the population, $\bar{L}(t)$, and all of it will be supplied regardless of its (rental) price—as long as this price is nonnegative. The labor market clearing condition can then be expressed as:

$$L(t) = \bar{L}(t) \tag{2.3}$$

for all t, where $L(t)$ denotes the demand for labor (and also the level of employment). More generally, this equation should be written in complementary slackness form. In particular, let the rental price of labor or the wage rate at time t be $w(t)$, then the labor market clearing condition takes the form

$$L(t) \leq \bar{L}(t), \; w(t) \geq 0 \quad \text{and} \quad \left(L(t) - \bar{L}(t)\right) w(t) = 0. \tag{2.4}$$

The complementary slackness formulation ensures that labor market clearing does not happen at a negative wage—or that if labor demand happens to be low enough, employment could be below $\bar{L}(t)$ at zero wage. However, this will not be an issue in most of the models studied in this book, because Assumption 1 and competitive labor markets ensure that wages are strictly positive (see Exercise 2.1). In view of this result, I use the simpler condition (2.3) throughout and denote both labor supply and employment at time t by $L(t)$.

The households also own the capital stock of the economy and rent it to firms. Let us denote the rental price of capital at time t by $R(t)$. The capital market clearing condition is similar to (2.3) and requires the demand for capital by firms to be equal to the supply of capital by households:

$$K(t) = \bar{K}(t),$$

where $\bar{K}(t)$ is the supply of capital by households and $K(t)$ is the demand by firms. Capital market clearing is straightforward to ensure in the class of models analyzed in this book. In particular, it is sufficient that the amount of capital $K(t)$ used in production at time t (from firms' optimization behavior) be consistent with households' endowments and saving behavior.

Let us take households' initial holdings of capital, $K(0) \geq 0$, as given (as part of the description of the environment). For now how this initial capital stock is distributed among the households is not important, since households' optimization decisions are not modeled explicitly and the economy is simply assumed to save a fraction s of its income. When we turn to models with household optimization below, an important part of the description of the environment will be to specify the preferences and the budget constraints of households.

At this point, I could also introduce the price of the final good at time t, say $P(t)$. But there is no need, since there is a choice of a numeraire commodity in this economy, whose price will be normalized to 1. In particular, as discussed in greater detail in Chapter 5, Walras's Law implies that the price of one of the commodities, the numeraire, should be normalized to 1. In fact, throughout I do something stronger and normalize the price of the final good to 1 in all periods. Ordinarily, one cannot choose more than one numeraire—otherwise, one would be fixing the relative price between the numeraires. But as explained in Chapter 5, we can build on an insight by Kenneth Arrow (Arrow, 1964) that it is sufficient to price *securities* (assets) that transfer one unit of consumption from one date (or state of the world) to another. In the context of dynamic economies, this implies that we need to keep track of an *interest rate* across periods, denoted by $r(t)$, which determines intertemporal prices and enables us to normalize the price of the final good to 1 within each period. Naturally we also need to keep track of the wage rate $w(t)$, which determines the price of labor relative to the final good at any date t.

This discussion highlights a central fact: all of the models in this book should be thought of as general equilibrium economies, in which different commodities correspond to the same good at different dates. Recall from basic general equilibrium theory that the same good at different dates (or in different states or localities) is a different commodity. Therefore, in almost all of the models in this book, there will be an infinite number of commodities, since time runs to infinity. This raises a number of special issues, which are discussed in Chapter 5 and later.

Returning to the basic Solow model, the next assumption is that capital depreciates, meaning that machines that are used in production lose some of their value because of wear and tear. In terms of the corn example above, some of the corn that is used as seed is no longer available for consumption or for use as seed in the following period. Let us assume that this depreciation takes an exponential form, which is mathematically very tractable. Thus capital depreciates (exponentially) at the rate $\delta \in (0, 1)$, so that out of 1 unit of capital this period, only $1 - \delta$ is left for next period. Though depreciation here stands for the wear and tear of the machinery, it can also represent the replacement of old machines by new ones in more realistic models (see Chapter 14).

The loss of part of the capital stock affects the interest rate (rate of return on savings) faced by households. Given the assumption of exponential depreciation at the rate δ and the normalization of the price of the final good to 1, the interest rate faced by the households is $r(t) = R(t) - \delta$, where recall that $R(t)$ is the rental price of capital at time t. A unit of final good can be consumed now or used as capital and rented to firms. In the latter case, a household receives $R(t)$ units of good in the next period as the rental price for its savings, but loses δ units of its capital holdings, since δ fraction of capital depreciates over time. Thus the household has given up one unit of commodity dated $t-1$ and receives $1 + r(t) = R(t) + 1 - \delta$ units of commodity dated t, so that $r(t) = R(t) - \delta$. The relationship between $r(t)$ and $R(t)$ explains the similarity between the symbols for the interest rate and the rental rate of capital. The interest rate faced by households plays a central role in the dynamic optimization decisions of households below. In the Solow model, this interest rate does not directly affect the allocation of resources.

2.1.3 Firm Optimization and Equilibrium

We are now in a position to look at the optimization problem of firms and the competitive equilibrium of this economy. Throughout the book I assume that the objective of firms is to maximize profits. Given the assumption that there is an aggregate production function, it is sufficient to consider the problem of a representative firm. Throughout, unless otherwise stated, I also assume that capital markets are functioning, so firms can rent capital in spot markets. For a given technology level $A(t)$, and given factor prices $R(t)$ and $w(t)$, the profit maximization problem of the representative firm at time t can be represented by the following static problem:

$$\max_{K\geq 0, L\geq 0} F(K, L, A(t)) - R(t)K - w(t)L. \tag{2.5}$$

When there are irreversible investments or costs of adjustments, as discussed, for example, in Section 7.8, the maximization problem of firms becomes dynamic. But in the absence of these features, maximizing profits separately at each date t is equivalent to maximizing the net present discounted value of profits. This feature simplifies the analysis considerably.

A couple of additional features are worth noting:

1. The maximization problem is set up in terms of aggregate variables, which, given the representative firm, is without any loss of generality.
2. There is nothing multiplying the F term, since the price of the final good has been normalized to 1. Thus the first term in (2.5) is the revenues of the representative firm (or the revenues of all of the firms in the economy).
3. This way of writing the problem already imposes competitive factor markets, since the firm is taking as given the rental prices of labor and capital, $w(t)$ and $R(t)$ (which are in terms of the numeraire, the final good).
4. This problem is concave, since F is concave (see Exercise 2.2).

An important aspect is that, because F exhibits constant returns to scale (Assumption 1), the maximization problem (2.5) does not have a well-defined solution (see Exercise 2.3); either there does not exist any (K, L) that achieves the maximum value of this program (which is infinity), or $K = L = 0$, or multiple values of (K, L) will achieve the maximum value of this program (when this value happens to be 0). This problem is related to the fact that in a world with constant returns to scale, the size of each individual firm is not determinate (only aggregates are determined). The same problem arises here because (2.5) is written without imposing the condition that factor markets should clear. A competitive equilibrium requires that all firms (and thus the representative firm) maximize profits and factor markets clear. In particular, the demands for labor and capital must be equal to the supplies of these factors at all times (unless the prices of these factors are equal to zero, which is ruled out by Assumption 1). This observation implies that the representative firm should make zero profits, since otherwise it would wish to hire arbitrarily large amounts of capital and labor exceeding the supplies, which are fixed. It also implies that total demand for labor, L, must be equal to the available supply of labor, $L(t)$. Similarly, the total demand for capital, K, should equal the total supply, $K(t)$. If this were not the case and $L < L(t)$, then there would be an excess supply of labor and the wage would be equal to zero. But this is not consistent with firm maximization, since given Assumption 1, the representative firm would then wish to hire an arbitrarily large amount of labor, exceeding the supply. This argument, combined with the fact that F is differentiable (Assumption 1), implies that given the supplies of capital and labor at time t, $K(t)$ and $L(t)$, factor prices must satisfy the following familiar conditions equating factor prices to marginal products:[1]

$$w(t) = F_L(K(t), L(t), A(t)), \tag{2.6}$$

and

$$R(t) = F_K(K(t), L(t), A(t)). \tag{2.7}$$

Euler's Theorem (Theorem 2.1) then verifies that at the prices (2.6) and (2.7), firms (or the representative firm) make zero profits.

Proposition 2.1 *Suppose Assumption 1 holds. Then, in the equilibrium of the Solow growth model, firms make no profits, and in particular,*

$$Y(t) = w(t)L(t) + R(t)K(t).$$

Proof. This result follows immediately from Theorem 2.1 for the case of constant returns to scale ($m = 1$). ■

Since firms make no profits in equilibrium, the ownership of firms does not need to be specified. All we need to know is that firms are profit-maximizing entities.

In addition to these standard assumptions on the production function, the following boundary conditions, the *Inada conditions,* are often imposed in the analysis of economic growth and macroeconomic equilibria.

Assumption 2 (Inada Conditions) *F satisfies the* Inada conditions

$$\lim_{K\to 0} F_K(K, L, A) = \infty \quad \text{and} \quad \lim_{K\to \infty} F_K(K, L, A) = 0 \text{ for all } L > 0 \text{ and all } A,$$

$$\lim_{L\to 0} F_L(K, L, A) = \infty \quad \text{and} \quad \lim_{L\to \infty} F_L(K, L, A) = 0 \text{ for all } K > 0 \text{ and all } A.$$

Moreover, $F(0, L, A) = 0$ for all L and A.

The role of these conditions—especially in ensuring the existence of *interior equilibria*—will become clear later in this chapter. They imply that the first units of capital and labor

1. An alternative way to derive (2.6) and (2.7) is to consider the cost minimization problem of the representative firm, which takes the form of minimizing $rK + wL$ with respect to K and L, subject to the constraint that $F(K, L, A) = Y$ for some level of output Y. This problem has a unique solution for any given level of Y. Then imposing market clearing, that is, $Y = F(K, L, A)$ with K and L corresponding to the supplies of capital and labor, yields (2.6) and (2.7).

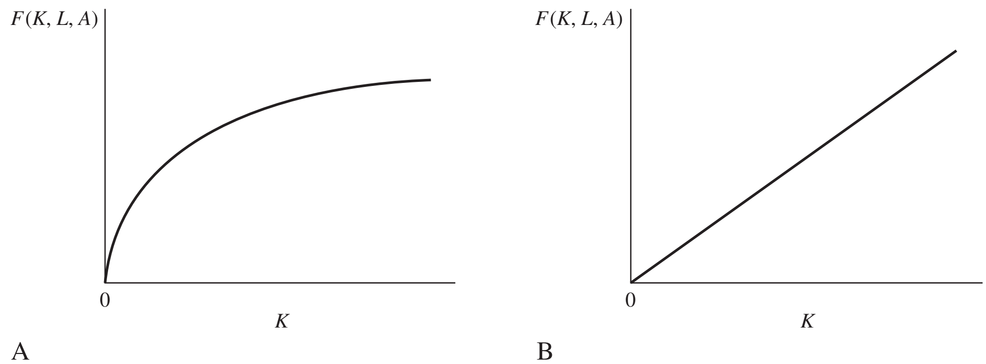

FIGURE 2.1 Production functions. (A) satisfies the Inada conditions in Assumption 2, while (B) does not.

are highly productive and that when capital or labor are sufficiently abundant, their marginal products are close to zero. The condition that $F(0, L, A) = 0$ for all L and A makes capital an essential input. This aspect of the assumption can be relaxed without any major implications for the results in this book. Figure 2.1 shows the production function $F(K, L, A)$ as a function of K, for given L and A, in two different cases; in panel A the Inada conditions are satisfied, while in panel B they are not.

I refer to Assumptions 1 and 2, which can be thought of as the neoclassical technology assumptions, throughout much of the book. For this reason, they are numbered independently from the equations, theorems, and proposition in this chapter.

2.2 The Solow Model in Discrete Time

I next present the dynamics of economic growth in the discrete-time Solow model.

2.2.1 Fundamental Law of Motion of the Solow Model

Recall that K depreciates exponentially at the rate δ, so that the law of motion of the capital stock is given by

$$K(t+1) = (1-\delta)\, K(t) + I(t), \tag{2.8}$$

where $I(t)$ is investment at time t.

From national income accounting for a closed economy, the total amount of final good in the economy must be either consumed or invested, thus

$$Y(t) = C(t) + I(t), \tag{2.9}$$

where $C(t)$ is consumption.[2] Using (2.1), (2.8), and (2.9), any feasible dynamic allocation in this economy must satisfy

$$K(t+1) \le F(K(t), L(t), A(t)) + (1-\delta)K(t) - C(t)$$

2. In addition, we can introduce government spending $G(t)$ on the right-hand side of (2.9). Government spending does not play a major role in the Solow growth model, thus its introduction is relegated to Exercise 2.7.

for $t = 0, 1, \ldots$. The question is to determine the equilibrium dynamic allocation among the set of feasible dynamic allocations. Here the behavioral rule that households save a constant fraction of their income simplifies the structure of equilibrium considerably (this is a behavioral rule, since it is not derived from the maximization of a well-defined utility function). One implication of this assumption is that any welfare comparisons based on the Solow model have to be taken with a grain of salt, since we do not know what the preferences of the households are.

Since the economy is closed (and there is no government spending), aggregate investment is equal to savings:

$$S(t) = I(t) = Y(t) - C(t).$$

The assumption that households save a constant fraction $s \in (0, 1)$ of their income can be expressed as

$$S(t) = sY(t), \tag{2.10}$$

which, in turn, implies that they consume the remaining $1 - s$ fraction of their income, and thus

$$C(t) = (1 - s)\, Y(t). \tag{2.11}$$

In terms of capital market clearing, (2.10) implies that the supply of capital for time $t + 1$ resulting from households' behavior can be expressed as $K(t + 1) = (1 - \delta)K(t) + S(t) = (1 - \delta)K(t) + sY(t)$. Setting supply and demand equal to each other and using (2.1) and (2.8) yields *the fundamental law of motion* of the Solow growth model:

$$K(t + 1) = sF(K(t), L(t), A(t)) + (1 - \delta)K(t). \tag{2.12}$$

This is a nonlinear difference equation. The equilibrium of the Solow growth model is described by (2.12) together with laws of motion for $L(t)$ and $A(t)$.

2.2.2 Definition of Equilibrium

The Solow model is a mixture of an old-style Keynesian model and a modern dynamic macroeconomic model. Households do not optimize when it comes to their savings or consumption decisions. Instead, their behavior is captured by (2.10) and (2.11). Nevertheless, firms still maximize profits, and factor markets clear. Thus it is useful to start defining equilibria in the way that is customary in modern dynamic macro models.

Definition 2.2 *In the basic Solow model for a given sequence of* $\{L(t), A(t)\}_{t=0}^{\infty}$ *and an initial capital stock* $K(0)$*, an* equilibrium path *is a sequence of capital stocks, output levels, consumption levels, wages, and rental rates* $\{K(t), Y(t), C(t), w(t), R(t)\}_{t=0}^{\infty}$ *such that* $K(t)$ *satisfies (2.12),* $Y(t)$ *is given by (2.1),* $C(t)$ *is given by (2.11), and* $w(t)$ *and* $R(t)$ *are given by (2.6) and (2.7), respectively.*

The most important point to note about Definition 2.2 is that an equilibrium is defined as an entire path of allocations and prices. An economic equilibrium does *not* refer to a static object; it specifies the entire path of behavior of the economy. Note also that Definition 2.2 incorporates the market clearing conditions, (2.6) and (2.7), into the definition of equilibrium. This practice is standard in macro and growth models. The alternative, which involves describing the equilibrium in more abstract terms, is discussed in Chapter 8 in the context of the neoclassical growth model (see, in particular, Definition 8.1).

2.2.3 Equilibrium without Population Growth and Technological Progress

It is useful to start with the following assumptions, which are relaxed later in this chapter:

1. There is no population growth; total population is constant at some level $L > 0$. Moreover, since households supply labor inelastically, this implies $L(t) = L$.
2. There is no technological progress, so that $A(t) = A$.

Let us define the capital-labor ratio of the economy as

$$k(t) \equiv \frac{K(t)}{L}, \tag{2.13}$$

which is a key object for the analysis. Now using the assumption of constant returns to scale, output (income) per capita, $y(t) \equiv Y(t)/L$, can be expressed as

$$\begin{aligned} y(t) &= F\left(\frac{K(t)}{L}, 1, A\right) \\ &\equiv f(k(t)). \end{aligned} \tag{2.14}$$

In other words, with constant returns to scale, output per capita is simply a function of the capital-labor ratio. Note that $f(k)$ here depends on A, so I could have written $f(k, A)$. I do not do this to simplify the notation and also because until Section 2.7, there will be no technological progress. Thus for now A is constant and can be normalized to $A = 1$.[3] The marginal product and the rental price of capital are then given by the derivative of F with respect to its first argument, which is $f'(k)$. The marginal product of labor and the wage rate are then obtained from Theorem 2.1, so that

$$\begin{aligned} R(t) &= f'(k(t)) > 0 \quad \text{and} \\ w(t) &= f(k(t)) - k(t)f'(k(t)) > 0. \end{aligned} \tag{2.15}$$

The fact that both factor prices are positive follows from Assumption 1, which ensures that the first derivatives of F with respect to capital and labor are always positive.

Example 2.1 (The Cobb-Douglas Production Function) *Let us consider the most common example of production function used in macroeconomics, the Cobb-Douglas production function. I hasten to add the caveat that even though the Cobb-Douglas form is convenient and widely used, it is also very special, and many interesting phenomena discussed later in this book are ruled out by this production function. The Cobb-Douglas production function can be written as*

$$\begin{aligned} Y(t) &= F(K(t), L(t), A(t)) \\ &= AK(t)^{\alpha}L(t)^{1-\alpha}, 0 < \alpha < 1. \end{aligned} \tag{2.16}$$

3. Later, when technological change is taken to be labor-augmenting, the term A can also be taken out, and the per capita production function can be written as $y = Af(k)$, with a slightly different definition of k as effective capital-labor ratio (see, e.g., (2.50) in Section 2.7).

It can easily be verified that this production function satisfies Assumptions 1 and 2, including the constant returns to scale feature imposed in Assumption 1. Dividing both sides by $L(t)$, the per capita production function in (2.14) becomes:

$$y(t) = Ak(t)^{\alpha},$$

where $y(t)$ again denotes output per worker and $k(t)$ is capital-labor ratio as defined in (2.13). The representation of factor prices as in (2.15) can also be verified. From the per capita production function representation, in particular (2.15), the rental price of capital can be expressed as

$$\begin{aligned} R(t) &= \frac{\partial Ak(t)^{\alpha}}{\partial k(t)}, \\ &= \alpha Ak(t)^{-(1-\alpha)}. \end{aligned}$$

Alternatively, in terms of the original production function (2.16), the rental price of capital in (2.7) is given by

$$\begin{aligned} R(t) &= \alpha AK(t)^{\alpha-1}L(t)^{1-\alpha} \\ &= \alpha Ak(t)^{-(1-\alpha)}, \end{aligned}$$

which is equal to the previous expression and thus verifies the form of the marginal product given in (2.15). Similarly, from (2.15),

$$\begin{aligned} w(t) &= Ak(t)^{\alpha} - \alpha Ak(t)^{-(1-\alpha)} \times k(t) \\ &= (1-\alpha)AK(t)^{\alpha}L(t)^{-\alpha}, \end{aligned}$$

which verifies the alternative expression for the wage rate in (2.6).

Returning to the analysis with the general production function, the per capita representation of the aggregate production function enables us to divide both sides of (2.12) by L to obtain the following simple difference equation for the evolution of the capital-labor ratio:

$$k(t+1) = sf(k(t)) + (1-\delta)k(t). \tag{2.17}$$

Since this difference equation is derived from (2.12), it also can be referred to as the *equilibrium difference equation* of the Solow model: it describes the equilibrium behavior of the key object of the model, the capital-labor ratio. The other equilibrium quantities can all be obtained from the capital-labor ratio $k(t)$.

At this point, let us also define a *steady-state equilibrium* for this model.

Definition 2.3 *A* steady-state equilibrium *without technological progress and population growth is an equilibrium path in which $k(t) = k^*$ for all t.*

In a steady-state equilibrium the capital-labor ratio remains constant. Since there is no population growth, this implies that the level of the capital stock will also remain constant. Mathematically, a steady-state equilibrium corresponds to a stationary point of the equilibrium difference equation (2.17). Most of the models in this book admit a steady-state equilibrium. This is also the case for this simple model.

The existence of a steady state can be seen by plotting the difference equation that governs the equilibrium behavior of this economy, (2.17), which is done in Figure 2.2. The thick curve

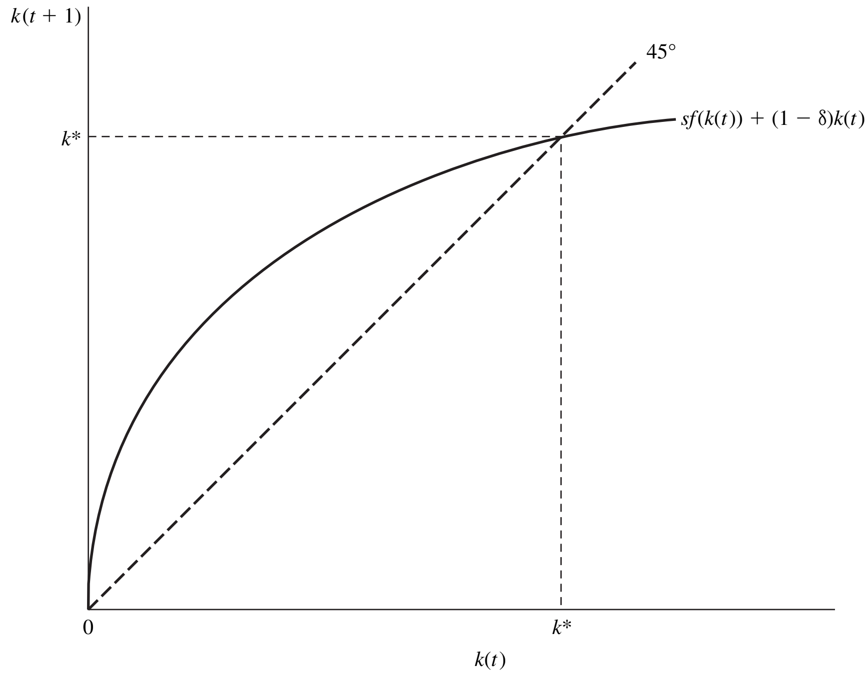

FIGURE 2.2 Determination of the steady-state capital-labor ratio in the Solow model without population growth and technological change.

represents the right-hand side of (2.17) and the dashed line corresponds to the 45° line. Their (positive) intersection gives the steady-state value of the capital-labor ratio k^*, which satisfies

$$\frac{f(k^*)}{k^*} = \frac{\delta}{s}. \tag{2.18}$$

Notice that in Figure 2.2 there is another intersection between (2.17) and the 45° line at $k = 0$. This second intersection occurs because, from Assumption 2, capital is an essential input, and thus $f(0) = 0$. Starting with $k(0) = 0$, there will then be no savings, and the economy will remain at $k = 0$. Nevertheless, I ignore this intersection throughout for a number of reasons. First, $k = 0$ is a steady-state equilibrium only when capital is an essential input and $f(0) = 0$. But as noted above, this assumption can be relaxed without any implications for the rest of the analysis, and when $f(0) > 0$, $k = 0$ is no longer a steady-state equilibrium. This is illustrated in Figure 2.3, which draws (2.17) for the case where $f(0) = \varepsilon$ for some $\varepsilon > 0$. Second, as we will see below, this intersection, even when it exists, is an unstable point; thus the economy would never travel toward this point starting with $K(0) > 0$ (or with $k(0) > 0$). Finally, and most importantly, this intersection holds no economic interest for us.[4]

An alternative visual representation shows the steady state as the intersection between a ray through the origin with slope δ (representing the function δk) and the function $sf(k)$. Figure 2.4, which illustrates this representation, is also useful for two other purposes. First, it depicts the levels of consumption and investment in a single figure. The vertical distance between the horizontal axis and the δk line at the steady-state equilibrium gives the amount of

4. Hakenes and Irmen (2006) show that even with $f(0) = 0$, the Inada conditions imply that in the continuous-time version of the Solow model $k = 0$ may not be the only equilibrium and the economy may move away from $k = 0$.

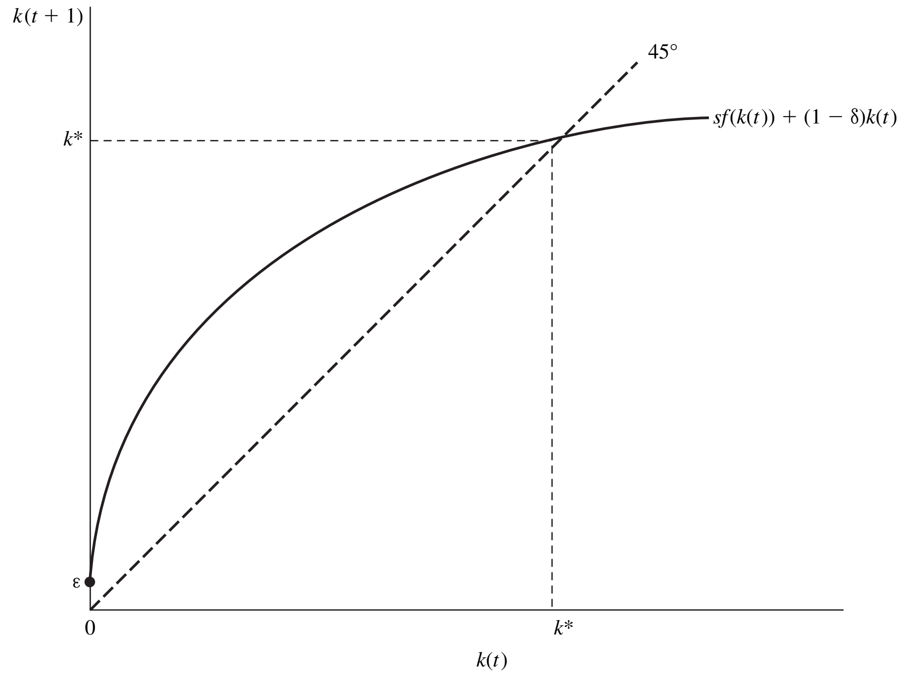

FIGURE 2.3 Unique steady state in the basic Solow model when $f(0) = \varepsilon > 0$.

investment per capita at the steady-state equilibrium (equal to δk^*), while the vertical distance between the function $f(k)$ and the δk line at k^* gives the level of consumption per capita. Clearly, the sum of these two terms make up $f(k^*)$. Second, Figure 2.4 also emphasizes that the steady-state equilibrium in the Solow model essentially sets investment, $sf(k)$, equal to the amount of capital that needs to be replenished, δk. This interpretation is particularly useful when population growth and technological change are incorporated.

This analysis therefore leads to the following proposition (with the convention that the intersection at $k = 0$ is being ignored even though $f(0) = 0$).

Proposition 2.2 *Consider the basic Solow growth model and suppose that Assumptions 1 and 2 hold. Then there exists a unique steady-state equilibrium where the capital-labor ratio* $k^* \in (0, \infty)$ *satisfies (2.18), per capita output is given by*

$$y^* = f(k^*), \tag{2.19}$$

and per capita consumption is given by

$$c^* = (1 - s)\, f(k^*). \tag{2.20}$$

Proof. The preceding argument establishes that any k^* that satisfies (2.18) is a steady state. To establish existence, note that from Assumption 2 (and from l'Hôpital's Rule, see Theorem A.21 in Appendix A), $\lim_{k\to 0} f(k)/k = \infty$ and $\lim_{k\to\infty} f(k)/k = 0$. Moreover, $f(k)/k$ is continuous from Assumption 1, so by the Intermediate Value Theorem (Theorem A.3) there exists k^* such that (2.18) is satisfied. To see uniqueness, differentiate $f(k)/k$ with respect to k, which gives

$$\frac{\partial(f(k)/k)}{\partial k} = \frac{f'(k)k - f(k)}{k^2} = -\frac{w}{k^2} < 0, \tag{2.21}$$

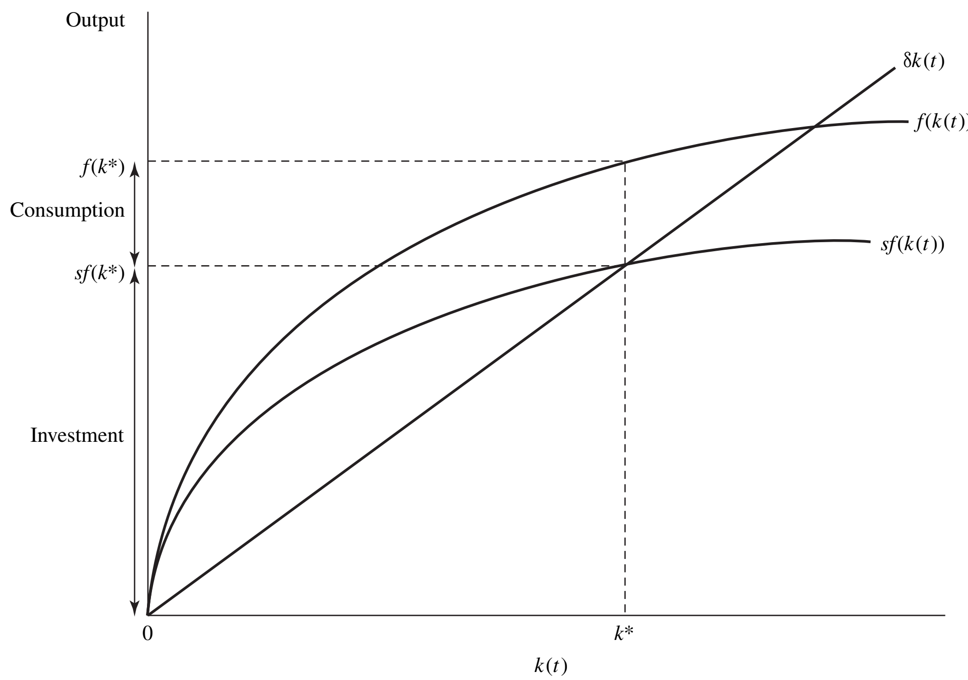

FIGURE 2.4 Investment and consumption in the steady-state equilibrium.

where the last equality in (2.21) uses (2.15). Since $f(k)/k$ is everywhere (strictly) decreasing, there can only exist a unique value k^* that satisfies (2.18). Equations (2.19) and (2.20) then follow by definition. ■

Through a series of examples, Figure 2.5 shows why Assumptions 1 and 2 cannot be dispensed with for establishing the existence and uniqueness results in Proposition 2.2. In the first two panels, the failure of Assumption 2 leads to a situation in which there is no steady-state equilibrium with positive activity, while in the third panel, the failure of Assumption 1 leads to nonuniqueness of steady states.

So far the model is very parsimonious: it does not have many parameters and abstracts from many features of the real world. An understanding of how cross-country differences in certain parameters translate into differences in growth rates or output levels is essential for our focus. This connection will be made in the next proposition. But before doing so, let us generalize the production function in one simple way and assume that

$$f(k) = A\tilde{f}(k),$$

where $A > 0$, so that A is a shift parameter, with greater values corresponding to greater productivity of factors. This type of productivity is referred to as "Hicks-neutral" (see below). For now, it is simply a convenient way of parameterizing productivity differences across countries. Since $f(k)$ satisfies the regularity conditions imposed above, so does $\tilde{f}(k)$.

Proposition 2.3 *Suppose Assumptions 1 and 2 hold and $f(k) = A\tilde{f}(k)$. Denote the steady-state level of the capital-labor ratio by $k^*(A, s, \delta)$ and the steady-state level of output by*

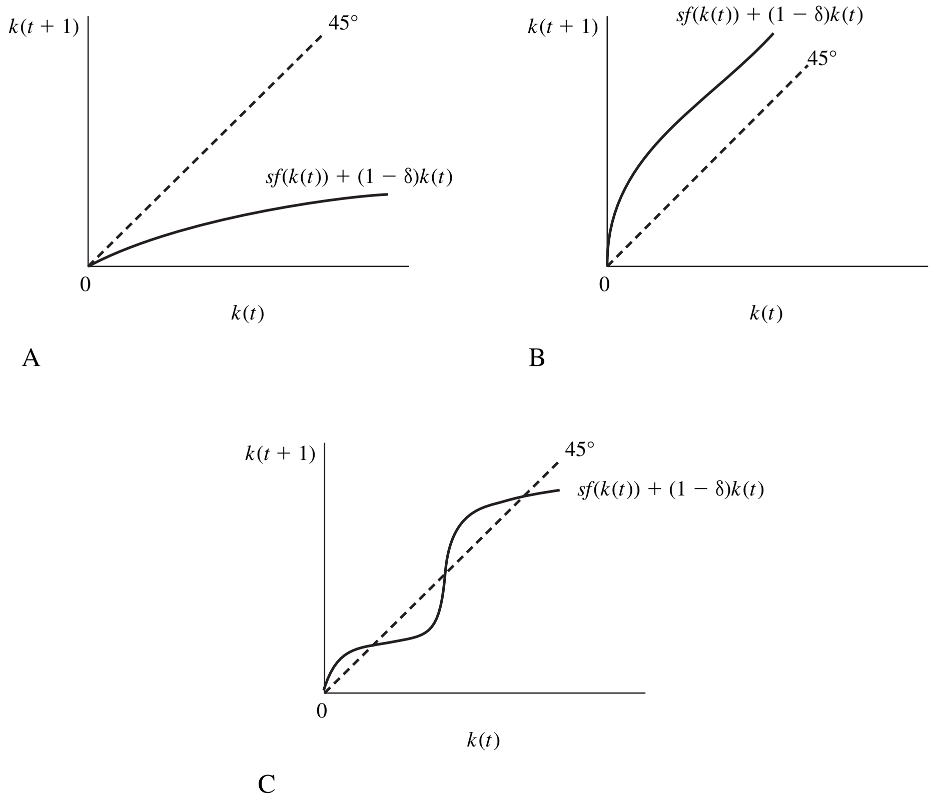

FIGURE 2.5 Examples of nonexistence and nonuniqueness of interior steady states when Assumptions 1 and 2 are not satisfied.

$y^*(A, s, \delta)$ when the underlying parameters are A, s, and δ. Then

$$\frac{\partial k^*(A, s, \delta)}{\partial A} > 0, \quad \frac{\partial k^*(A, s, \delta)}{\partial s} > 0, \quad \text{and } \frac{\partial k^*(A, s, \delta)}{\partial \delta} < 0;$$
$$\frac{\partial y^*(A, s, \delta)}{\partial A} > 0, \quad \frac{\partial y^*(A, s, \delta)}{\partial s} > 0, \quad \text{and } \frac{\partial y^*(A, s, \delta)}{\partial \delta} < 0.$$

Proof. The proof follows immediately by writing

$$\frac{\tilde{f}(k^*)}{k^*} = \frac{\delta}{As},$$

which holds for an open set of values of k^*, A, s, and δ. Now apply the Implicit Function Theorem (Theorem A.25) to obtain the results. For example,

$$\frac{\partial k^*}{\partial s} = \frac{\delta (k^*)^2}{s^2 w^*} > 0,$$

where $w^* = f(k^*) - k^* f'(k^*) > 0$. The other results follow similarly. ■

Therefore countries with higher saving rates and better technologies will have higher capital-labor ratios and will be richer. Those with greater (technological) depreciation will tend to have lower capital-labor ratios and will be poorer. All of the results in Proposition 2.3 are intuitive, and they provide us with a first glimpse of the potential determinants of the capital-labor ratios and output levels across countries.

The same comparative statics with respect to A and δ also apply to c^*. However, it is straightforward to see that c^* is not monotone in the saving rate (e.g., think of the extreme case where $s = 1$). In fact, there exists a unique saving rate, s_{gold}, referred to as the "golden rule" saving rate, which maximizes the steady-state level of consumption. Since we are treating the saving rate as an exogenous parameter and have not specified the objective function of households yet, we cannot say whether the golden rule saving rate is better than some other saving rate. It is nevertheless interesting to characterize what this golden rule saving rate corresponds to. To do this, let us first write the steady-state relationship between c^* and s and suppress the other parameters:

$$\begin{aligned} c^*(s) &= (1-s)f(k^*(s)) \\ &= f(k^*(s)) - \delta k^*(s), \end{aligned}$$

where the second equality exploits the fact that in steady state, $sf(k) = \delta k$. Now differentiating this second line with respect to s (again using the Implicit Function Theorem), we obtain

$$\frac{\partial c^*(s)}{\partial s} = [f'(k^*(s)) - \delta]\frac{\partial k^*}{\partial s}. \tag{2.22}$$

Let us define the golden rule saving rate s_{gold} to be such that $\partial c^*(s_{\text{gold}})/\partial s = 0$. The corresponding steady-state golden rule capital stock is defined as k^*_{gold}. These quantities and the relationship between consumption and the saving rate are plotted in Figure 2.6. The next proposition shows that s_{gold} and k^*_{gold} are uniquely defined.

Proposition 2.4 *In the basic Solow growth model, the highest level of steady-state consumption is reached for s_{gold}, with the corresponding steady-state capital level k^*_{gold} such that*

$$f'(k^*_{\text{gold}}) = \delta. \tag{2.23}$$

Proof. By definition $\partial c^*(s_{\text{gold}})/\partial s = 0$. From Proposition 2.3, $\partial k^*/\partial s > 0$; thus (2.22) can be equal to zero only when $f'(k^*(s_{\text{gold}})) = \delta$. Moreover, when $f'(k^*(s_{\text{gold}})) = \delta$, it can be verified that $\partial^2 c^*(s_{\text{gold}})/\partial s^2 < 0$, so $f'(k^*(s_{\text{gold}})) = \delta$ indeed corresponds to a local maximum. That $f'(k^*(s_{\text{gold}})) = \delta$ also yields the global maximum is a consequence of the following observations: for all $s \in [0, 1]$, we have $\partial k^*/\partial s > 0$, and moreover, when $s < s_{\text{gold}}$, $f'(k^*(s)) - \delta > 0$ by the concavity of f, so $\partial c^*(s)/\partial s > 0$ for all $s < s_{\text{gold}}$. By the converse argument, $\partial c^*(s)/\partial s < 0$ for all $s > s_{\text{gold}}$. Therefore only s_{gold} satisfies $f'(k^*(s)) = \delta$ and gives the unique global maximum of consumption per capita. ■

In other words, there exists a unique saving rate, s_{gold}, and also a unique corresponding capital-labor ratio, k^*_{gold}, given by (2.23), that maximize the level of steady-state consumption. When the economy is below k^*_{gold}, a higher saving rate will increase consumption, whereas when the economy is above k^*_{gold}, steady-state consumption can be raised by saving less. In the latter case, lower savings translate into higher consumption, because the capital-labor ratio of the economy is too high; households are investing too much and not consuming enough. This is the essence of the phenomenon of *dynamic inefficiency,* discussed in greater detail in Chapter 9. For now, recall that there is no explicit utility function here, so statements about inefficiency

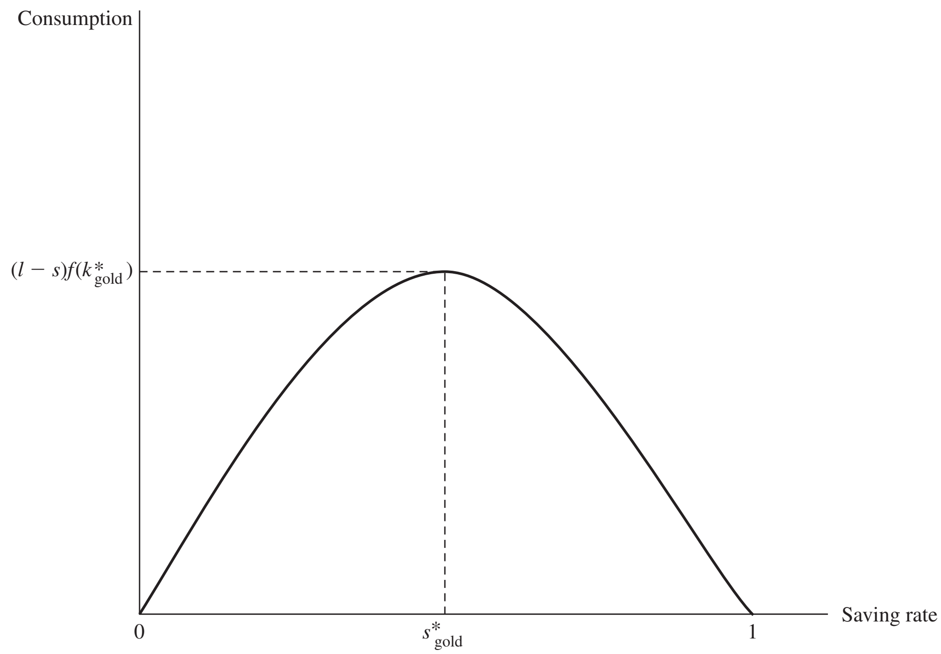

FIGURE 2.6 The golden rule level of saving rate, which maximizes steady-state consumption.

must be considered with caution. In fact, the reason this type of dynamic inefficiency does not generally apply when consumption-saving decisions are endogenized may already be apparent to many of you.

2.3 Transitional Dynamics in the Discrete-Time Solow Model

Proposition 2.2 establishes the existence of a unique steady-state equilibrium (with positive activity). Recall that an equilibrium path does not refer simply to the steady state but to the entire path of capital stock, output, consumption, and factor prices. This is an important point to bear in mind, especially since the term "equilibrium" is used differently in economics than in other disciplines. Typically, in engineering and the physical sciences, an equilibrium refers to a point of rest of a dynamical system, thus to what I have so far referred to as "the steady-state equilibrium." One may then be tempted to say that the system is in "disequilibrium" when it is away from the steady state. However, in economics, the non-steady-state behavior of an economy is also governed by market clearing and optimizing behavior of households and firms. Most economies spend much of their time in non-steady-state situations. Thus we are typically interested in the entire dynamic equilibrium path of the economy, not just in its steady state.

To determine what the equilibrium path of our simple economy looks like, we need to study the transitional dynamics of the equilibrium difference equation (2.17) starting from an arbitrary capital-labor ratio, $k(0) > 0$. Of special interest are the answers to the questions of whether the economy will tend to this steady state starting from an arbitrary capital-labor ratio and how it will behave along the transition path. Recall that the total amount of capital at the beginning of the economy, $K(0) > 0$, is taken as a state variable, while for now, the supply of labor L is fixed. Therefore at time $t = 0$, the economy starts with an arbitrary capital-labor ratio $k(0) = K(0)/L > 0$ as its initial value and then follows the law of motion given by the difference equation (2.17). Thus the question is whether (2.17) will take us to the unique steady state starting from an arbitrary initial capital-labor ratio.

Before answering this question, recall some definitions and key results from the theory of dynamical systems. Appendix B provides more details and a number of further results. Consider the nonlinear system of autonomous difference equations,

$$\mathbf{x}(t+1) = \mathbf{G}(\mathbf{x}(t)), \tag{2.24}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$ (where $n \in \mathbb{R}$). Let $\mathbf{x}^*$ be a *fixed point* of the mapping $\mathbf{G}(\cdot)$, that is,

$$\mathbf{x}^* = \mathbf{G}(\mathbf{x}^*).$$

I refer to $\mathbf{x}^*$ as a "steady state" of the difference equation (2.24).[5] The relevant notion of stability is introduced in the next definition.

Definition 2.4 *A steady state* $\mathbf{x}^*$ *is* locally asymptotically stable *if there exists an open set* $B(\mathbf{x}^*)$ *containing* $\mathbf{x}^*$ *such that for any solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$ *to (2.24) with* $\mathbf{x}(0) \in B(\mathbf{x}^*)$, $\mathbf{x}(t) \to \mathbf{x}^*$. *Moreover,* $\mathbf{x}^*$ *is* globally asymptotically stable *if for all* $\mathbf{x}(0) \in \mathbb{R}^n$, *for any solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$, $\mathbf{x}(t) \to \mathbf{x}^*$.

The next theorem provides the main results on the stability properties of systems of linear difference equations. The following theorems are special cases of the results presented in Appendix B.

Theorem 2.2 (Stability for Systems of Linear Difference Equations) *Consider the following linear difference equation system:*

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}, \tag{2.25}$$

with initial value $\mathbf{x}(0)$*, where* $\mathbf{x}(t) \in \mathbb{R}^n$ *for all* t*,* $\mathbf{A}$ *is an* $n \times n$ *matrix, and* $\mathbf{b}$ *is a* $n \times 1$ *column vector. Let* $\mathbf{x}^*$ *be the steady state of the difference equation given by* $\mathbf{A}\mathbf{x}^* + \mathbf{b} = \mathbf{x}^*$*. Suppose that all of the eigenvalues of* $\mathbf{A}$ *are strictly inside the unit circle in the complex plane. Then the steady state of the difference equation (2.25),* $\mathbf{x}^*$*, is globally (asymptotically) stable, in the sense that starting from any* $\mathbf{x}(0) \in \mathbb{R}^n$*, the unique solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$ *satisfies* $\mathbf{x}(t) \to \mathbf{x}^*$.

Unfortunately, much less can be said about nonlinear systems, but the following is a standard local stability result.

Theorem 2.3 (Local Stability for Systems of Nonlinear Difference Equations) *Consider the following nonlinear autonomous system:*

$$\mathbf{x}(t+1) = \mathbf{G}(\mathbf{x}(t)), \tag{2.26}$$

with initial value $\mathbf{x}(0)$*, where* $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$*. Let* $\mathbf{x}^*$ *be a steady state of this system, that is,* $\mathbf{G}(\mathbf{x}^*) = \mathbf{x}^*$*, and suppose that* $\mathbf{G}$ *is differentiable at* $\mathbf{x}^*$*. Define*

$$\mathbf{A} \equiv D\mathbf{G}(\mathbf{x}^*),$$

5. Various other terms are used to describe $\mathbf{x}^*$, for example, "equilibrium point" or "critical point." Since these other terms have different meanings in economics, I refer to $\mathbf{x}^*$ as a steady state throughout.

where $D\mathbf{G}$ denotes the matrix of partial derivatives (Jacobian) of $\mathbf{G}$. Suppose that all of the eigenvalues of $\mathbf{A}$ are strictly inside the unit circle. Then the steady state of the difference equation (2.26), $\mathbf{x}^*$, is locally (asymptotically) stable, in the sense that there exists an open neighborhood of $\mathbf{x}^*$, $\mathbf{B}(\mathbf{x}^*) \subset \mathbb{R}^n$, such that starting from any $\mathbf{x}(0) \in \mathbf{B}(\mathbf{x}^*)$, $\mathbf{x}(t) \to \mathbf{x}^*$.

An immediate corollary of Theorem 2.3 is the following useful result.

Corollary 2.1

1. *Let $x(t), a, b \in \mathbb{R}$. If $|a| < 1$, then the unique steady state of the linear difference equation $x(t+1) = ax(t) + b$ is globally (asymptotically) stable, in the sense that $x(t) \to x^* = b/(1-a)$.*
2. *Let $g : \mathbb{R} \to \mathbb{R}$ be differentiable in the neighborhood of the steady state x^*, defined by $g(x^*) = x^*$, and suppose that $\left|g'(x^*)\right| < 1$. Then the steady state x^* of the nonlinear difference equation $x(t+1) = g(x(t))$ is locally (asymptotically) stable. Moreover, if g is continuously differentiable and satisfies $\left|g'(x)\right| < 1$ for all $x \in \mathbb{R}$, then x^* is globally (asymptotically) stable.*

Proof. The first part follows immediately from Theorem 2.2. The local stability of g in the second part follows from Theorem 2.3. Global stability follows since

$$\begin{aligned}|x(t+1) - x^*| &= |g(x(t)) - g(x^*)| \\ &= \left| \int_{x^*}^{x(t)} g'(x)dx \right| \\ &< |x(t) - x^*|,\end{aligned}$$

where the second line follows from the Fundamental Theorem of Calculus (Theorem B.2 in Appendix B), and the last inequality uses the hypothesis that $\left|g'(x)\right| < 1$ for all $x \in \mathbb{R}$. This implies that for any $x(0) < x^*$, $\{x(t)\}_{t=0}^{\infty}$ is an increasing sequence. Since $\left|g'(x)\right| < 1$, there cannot exist $x' \neq x^*$ such that $x' = g(x')$, and moreover $\{x(t)\}_{t=0}^{\infty}$ is bounded above by x^*. It therefore converges to x^*. The argument for the case where $x(0) > x^*$ is identical. ■

We can now apply Corollary 2.1 to the equilibrium difference equation (2.17) of the Solow model to establish the local stability of the steady-state equilibrium. Global stability does not directly follow form Corollary 2.1 (since the equivalent of $\left|g'(x)\right| < 1$ for all x is not true), but a slightly different argument can be used to prove this property.

Proposition 2.5 *Suppose that Assumptions 1 and 2 hold. Then the steady-state equilibrium of the Solow growth model described by the difference equation (2.17) is globally asymptotically stable, and starting from any $k(0) > 0$, $k(t)$ monotonically converges to k^*.*

Proof. Let $g(k) \equiv sf(k) + (1-\delta)k$. First observe that $g'(k)$ exists and is always strictly positive, that is, $g'(k) > 0$ for all k. Next, from (2.17),

$$k(t+1) = g(k(t)), \tag{2.27}$$

with a unique steady state at k^*. From (2.18), the steady-state capital k^* satisfies $\delta k^* = sf(k^*)$, or

$$k^* = g(k^*). \tag{2.28}$$

Now recall that $f(\cdot)$ is concave and differentiable from Assumption 1 and satisfies $f(0) = 0$ from Assumption 2. For any strictly concave differentiable function, we have (recall Fact A.23 in Appendix A):

$$f(k) > f(0) + kf'(k) = kf'(k). \tag{2.29}$$

Since (2.29) implies that $\delta = sf(k^*)/k^* > sf'(k^*)$, we have $g'(k^*) = sf'(k^*) + 1 - \delta < 1$. Therefore

$$g'(k^*) \in (0, 1).$$

Corollary 2.1 then establishes local asymptotic stability.

To prove global stability, note that for all $k(t) \in (0, k^*)$,

$$\begin{aligned} k(t+1) - k^* &= g(k(t)) - g(k^*) \\ &= -\int_{k(t)}^{k^*} g'(k)dk, \\ &< 0, \end{aligned}$$

where the first line follows by subtracting (2.28) from (2.27), the second line again uses the Fundamental Theorem of Calculus (Theorem B.2), and the third line follows from the observation that $g'(k) > 0$ for all k. Next, (2.17) also implies

$$\begin{aligned} \frac{k(t+1) - k(t)}{k(t)} &= s\frac{f(k(t))}{k(t)} - \delta \\ &> s\frac{f(k^*)}{k^*} - \delta \\ &= 0, \end{aligned}$$

where the second line uses the fact that $f(k)/k$ is decreasing in k (from (2.29)) and the last line uses the definition of k^*. These two arguments together establish that for all $k(t) \in (0, k^*)$, $k(t+1) \in (k(t), k^*)$. Therefore $\{k(t)\}_{t=0}^{\infty}$ is monotonically increasing and is bounded above by k^*. Moreover, since k^* is the unique steady state (with $k > 0$), there exists no $k' \in (0, k^*)$ such that $k(t+1) = k(t) = k'$ for any t. Therefore $\{k(t)\}_{t=0}^{\infty}$ must monotonically converge to k^*. An identical argument implies that for all $k(t) > k^*$, $k(t+1) \in (k^*, k(t))$ and establishes monotonic convergence starting from $k(0) > k^*$. This completes the proof of global stability. ■

This stability result can be seen diagrammatically in Figure 2.7. Starting from initial capital stock $k(0) > 0$, which is below the steady-state level k^*, the economy grows toward k^* and experiences *capital deepening*—meaning that the capital-labor ratio increases. Together with capital deepening comes growth of per capita income. If instead the economy were to start with $k'(0) > k^*$, it would reach the steady state by decumulating capital and contracting (i.e., by experiencing negative growth).

The following proposition is an immediate corollary of Proposition 2.5.

Proposition 2.6 *Suppose that Assumptions 1 and 2 hold, and $k(0) < k^*$. Then $\{w(t)\}_{t=0}^{\infty}$ is an increasing sequence, and $\{R(t)\}_{t=0}^{\infty}$ is a decreasing sequence. If $k(0) > k^*$, the opposite results apply.*

Proof. See Exercise 2.9. ■

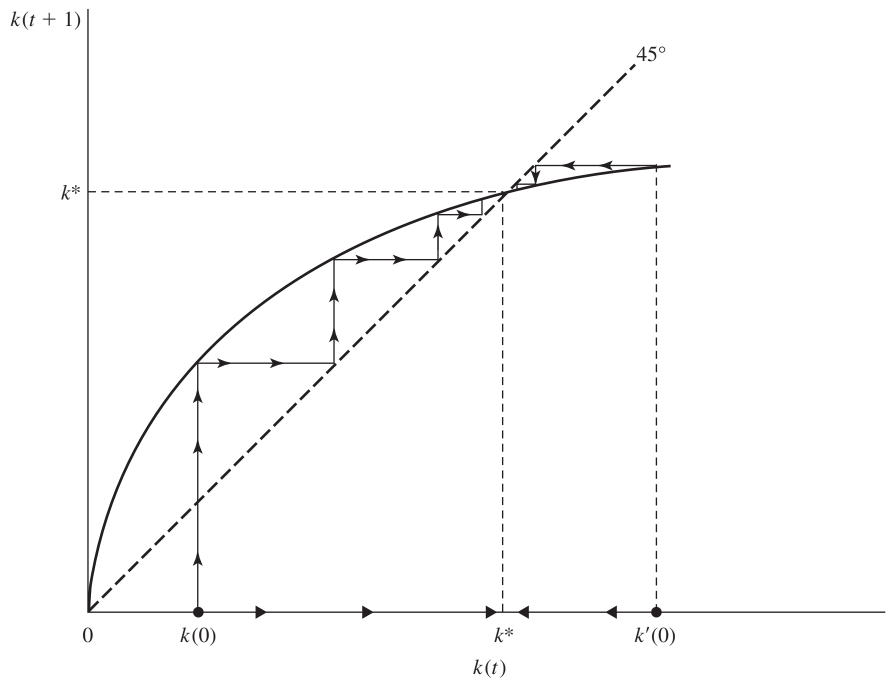

FIGURE 2.7 Transitional dynamics in the basic Solow model.

Recall that when the economy starts with too little capital relative to its labor supply, the capital-labor ratio will increase. Thus the marginal product of capital will fall due to diminishing returns to capital and the wage rate will increase. Conversely, if it starts with too much capital, it will decumulate capital, and in the process the wage rate will decline and the rate of return to capital will increase.

The analysis has established that the Solow growth model has a number of nice properties: unique steady state, global (asymptotic) stability, and finally, simple and intuitive comparative statics. Yet so far it has no growth. The steady state is the point at which there is no growth in the capital-labor ratio, no more capital deepening, and no growth in output per capita. Consequently, the basic Solow model (without technological progress) can only generate economic growth along the transition path to the steady state (starting with $k(0) < k^*$). However this growth is not sustained: it slows down over time and eventually comes to an end. Section 2.7 shows that the Solow model can incorporate economic growth by allowing exogenous technological change. Before doing this, it is useful to look at the relationship between the discrete- and continuous-time formulations.

2.4 The Solow Model in Continuous Time

2.4.1 From Difference to Differential Equations

Recall that the time periods $t = 0, 1, \ldots$ can refer to days, weeks, months, or years. In some sense, the time unit is not important. This arbitrariness suggests that perhaps it may be more convenient to look at dynamics by making the time unit as small as possible, that is, by going to continuous time. While much of modern macroeconomics (outside of growth theory) uses discrete-time models, many growth models are formulated in continuous time. The continuous-time setup has a number of advantages, since some pathological results of discrete-time models disappear when using continuous time (see Exercise 2.21). Moreover, continuous-time models have more flexibility in the analysis of dynamics and allow explicit-form solutions in a wider set of circumstances. These considerations motivate the detailed study of both the discrete- and continuous-time versions of the basic models in this book.

Let us start with a simple difference equation:

$$x(t+1) - x(t) = g(x(t)). \tag{2.30}$$

This equation states that between time t and $t+1$, the absolute growth in x is given by $g(x(t))$. Imagine that time is more finely divisible than that captured by our discrete indices, $t = 0, 1, \ldots$. In the limit, we can think of time as being as finely divisible as we would like, so that $t \in \mathbb{R}_+$. In that case, (2.30) gives us information about how the variable x changes between two discrete points in time, t and $t+1$. Between these time periods, we do not know how x evolves. However, if t and $t+1$ are not too far apart, the following approximation is reasonable:

$$x(t + \Delta t) - x(t) \simeq \Delta t \cdot g(x(t))$$

for any $\Delta t \in [0, 1]$. When $\Delta t = 0$, this equation is just an identity. When $\Delta t = 1$, it gives (2.30). In between it is a linear approximation. This approximation will be relatively accurate if the distance between t and $t+1$ is not very large, so that $g(x) \simeq g(x(t))$ for all $x \in [x(t), x(t+1)]$ (however, you should also convince yourself that this approximation could in fact be quite bad if the function g is highly nonlinear, in which case its behavior changes significantly between $x(t)$ and $x(t+1)$). Now divide both sides of this equation by Δt, and take limits to obtain

$$\lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = \dot{x}(t) \simeq g(x(t)), \tag{2.31}$$

where, as throughout the book, I use the dot notation to denote time derivatives, $\dot{x}(t) \equiv dx(t)/dt$. Equation (2.31) is a differential equation representing the same dynamics as the difference equation (2.30) for the case in which the distance between t and $t+1$ is small.

2.4.2 The Fundamental Equation of the Solow Model in Continuous Time

We can now repeat all of the analysis so far using the continuous-time representation. Nothing has changed on the production side, so we continue to have (2.6) and (2.7) as the factor prices, but now these refer to instantaneous rental rates. For example, $w(t)$ is the flow of wages that workers receive at instant t. Savings are again given by

$$S(t) = sY(t),$$

while consumption is still given by (2.11).

Let us also introduce population growth into this model and assume that the labor force $L(t)$ grows proportionally, that is,

$$L(t) = \exp(nt)L(0). \tag{2.32}$$

The purpose of doing so is that in many of the classical analyses of economic growth, population growth plays an important role, so it is useful to see how it affects the equilibrium here. There is still no technological progress.

Recall that

$$k(t) \equiv \frac{K(t)}{L(t)},$$

which implies that

$$\frac{\dot{k}(t)}{k(t)} = \frac{\dot{K}(t)}{K(t)} - \frac{\dot{L}(t)}{L(t)},$$
$$= \frac{\dot{K}(t)}{K(t)} - n,$$

where I used the fact that, from (2.32), $\dot{L}(t)/L(t) = n$. From the limiting argument leading to equation (2.31) in the previous subsection, the law of motion of the capital stock is given by

$$\dot{K}(t) = sF(K(t), L(t), A(t)) - \delta K(t).$$

Using the definition of $k(t)$ as the capital-labor ratio and the constant returns to scale properties of the production function, the fundamental law of motion of the Solow model in continuous time is obtained as

$$\frac{\dot{k}(t)}{k(t)} = s\frac{f(k(t))}{k(t)} - (n + \delta), \tag{2.33}$$

where, following usual practice, I have transformed the left-hand side to the proportional change in the capital-labor ratio by dividing both sides by $k(t)$.[6]

Definition 2.5 *In the basic Solow model in continuous time with population growth at the rate n, no technological progress and an initial capital stock $K(0)$, an* equilibrium path *is given by paths (sequences) of capital stocks, labor, output levels, consumption levels, wages, and rental rates $[K(t), L(t), Y(t), C(t), w(t), R(t)]_{t=0}^{\infty}$ such that $L(t)$ satisfies (2.32), $k(t) \equiv K(t)/L(t)$ satisfies (2.33), $Y(t)$ is given by (2.1), $C(t)$ is given by (2.11), and $w(t)$ and $R(t)$ are given by (2.6) and (2.7), respectively.*

As before, a steady-state equilibrium involves $k(t)$ remaining constant at some level k^*.

It is easy to verify that the equilibrium differential equation (2.33) has a unique steady state at k^*, which is given by a slight modification of (2.18) to incorporate population growth:

$$\frac{f(k^*)}{k^*} = \frac{n + \delta}{s}. \tag{2.34}$$

In other words, going from discrete to continuous time has not changed any of the basic economic features of the model. Thus the steady state can again be plotted in a diagram similar to Figure 2.1 except that it now also incorporates population growth. This is done in Figure 2.8, which also highlights that the logic of the steady state is the same with population growth as it was without population growth. The amount of investment, $sf(k)$, is used to replenish the capital-labor ratio, but now there are two reasons for replenishments. The capital stock depreciates exponentially at the flow rate δ. In addition, the capital stock must also increase as

6. Throughout I adopt the notation $[x(t)]_{t=0}^{\infty}$ to denote the continuous-time path of variable $x(t)$. An alternative notation often used in the literature is $(x(t); t \geq 0)$. I prefer the former both because it is slightly more compact and also because it is more similar to the discrete-time notation for the time path of a variable, $\{x(t)\}_{t=0}^{\infty}$. When referring to $[x(t)]_{t=0}^{\infty}$, I use the terms "path," "sequence," and "function (of time t)" interchangeably.

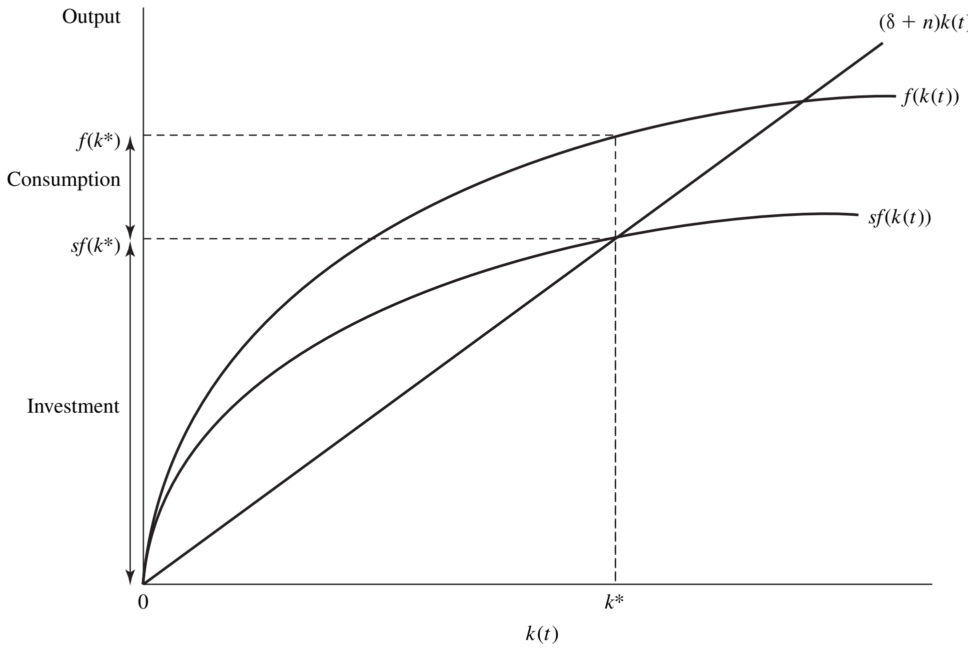

FIGURE 2.8 Investment and consumption in the steady-state equilibrium with population growth.

population grows to maintain the capital-labor ratio at a constant level. The amount of capital that needs to be replenished is therefore $(n + \delta)k$.

Proposition 2.7 *Consider the basic Solow growth model in continuous time and suppose that Assumptions 1 and 2 hold. Then there exists a unique steady-state equilibrium where the capital-labor ratio is equal to $k^* \in (0, \infty)$ and satisfies (2.34), per capita output is given by*

$$y^* = f(k^*),$$

and per capita consumption is given by

$$c^* = (1 - s)f(k^*).$$

Proof. See Exercise 2.5. ■

Moreover, again defining $f(k) = A\tilde{f}(k)$, the following proposition holds.

Proposition 2.8 *Suppose Assumptions 1 and 2 hold and $f(k) = A\tilde{f}(k)$. Denote the steady-state equilibrium level of the capital-labor ratio by $k^*(A, s, \delta, n)$ and the steady-state level of output by $y^*(A, s, \delta, n)$ when the underlying parameters are given by A, s, δ, and n. Then we have*

$$\frac{\partial k^*(A,s,\delta,n)}{\partial A} > 0, \quad \frac{\partial k^*(A,s,\delta,n)}{\partial s} > 0, \quad \frac{\partial k^*(A,s,\delta,n)}{\partial \delta} < 0, \quad \text{and} \quad \frac{\partial k^*(A,s,\delta,n)}{\partial n} < 0;$$
$$\frac{\partial y^*(A,s,\delta,n)}{\partial A} > 0, \quad \frac{\partial y^*(A,s,\delta,n)}{\partial s} > 0, \quad \frac{\partial y^*(A,s,\delta,n)}{\partial \delta} < 0, \quad \text{and} \quad \frac{\partial y^*(A,s,\delta,n)}{\partial n} < 0.$$

Proof. See Exercise 2.6. ■

The new result relative to the earlier comparative static proposition (Proposition 2.3) is that now a higher population growth rate, n, also reduces the capital-labor ratio and output per

2

The Solow Growth Model

The previous chapter introduced a number of basic facts and posed the main questions concerning the sources of economic growth over time and the causes of differences in economic performance across countries. These questions are central not only to growth theory but also to macroeconomics and the social sciences more generally. Our next task is to develop a simple framework that can help us think about the proximate causes and the mechanics of the process of economic growth and cross-country income differences. We will use this framework both to study potential sources of economic growth and also to perform simple comparative statics to gain an understanding of which country characteristics are conducive to higher levels of income per capita and more rapid economic growth.

Our starting point is the so-called Solow-Swan model named after Robert (Bob) Solow and Trevor Swan, or simply the Solow model, named after the more famous of the two economists. These economists published two pathbreaking articles in the same year, 1956 (Solow, 1956; Swan, 1956) introducing the Solow model. Bob Solow later developed many implications and applications of this model and was awarded the Nobel prize in economics for his contributions. This model has shaped the way we approach not only economic growth but also the entire field of macroeconomics. Consequently, a by-product of our analysis of this chapter is a detailed exposition of a workhorse model of macroeconomics.

The Solow model is remarkable in its simplicity. Looking at it today, one may fail to appreciate how much of an intellectual breakthrough it was. Before the advent of the Solow growth model, the most common approach to economic growth built on the model developed by Roy Harrod and Evsey Domar (Harrod, 1939; Domar, 1946). The Harrod-Domar model emphasized potential dysfunctional aspects of economic growth, for example, how economic growth could go hand-in-hand with increasing unemployment (see Exercise 2.23 on this model). The Solow model demonstrated why the Harrod-Domar model was not an attractive place to start. At the center of the Solow growth model, distinguishing it from the Harrod-Domar model, is the neoclassical aggregate production function. This function not only enables the Solow model to make contact with microeconomics, but as we will see in the next chapter, it also serves as a bridge between the model and the data.

An important feature of the Solow model, which is shared by many models presented in this book, is that it is a simple and abstract representation of a complex economy. At first, it may appear too simple or too abstract. After all, to do justice to the process of growth or macroeconomic equilibrium, we have to consider households and individuals with different tastes, abilities, incomes, and roles in society; various sectors; and multiple social interactions. The Solow model cuts through these complications by constructing a simple one-good economy, with little reference to individual decisions. Therefore, the Solow model should be thought of as a starting point and a springboard for richer models.

In this chapter, I present the basic Solow model. The closely related neoclassical growth model is presented in Chapter 8.

2.1 The Economic Environment of the Basic Solow Model

Economic growth and development are dynamic processes and thus necessitate dynamic models. Despite its simplicity, the Solow growth model is a dynamic general equilibrium model (though, importantly, many key features of dynamic general equilibrium models emphasized in Chapter 5, such as preferences and dynamic optimization, are missing in this model).

The Solow model can be formulated in either discrete or continuous time. I start with the discrete-time version, because it is conceptually simpler and more commonly used in macroeconomic applications. However, many growth models are formulated in continuous time, and I then provide a detailed exposition of the continuous-time version of the Solow model and show that it is often more convenient to work with.

2.1.1 Households and Production

Consider a closed economy, with a unique final good. The economy is in discrete time running to an infinite horizon, so that time is indexed by $t = 0, 1, 2, \ldots$. Time periods here may correspond to days, weeks, or years. For now, we do not need to specify the time scale.

The economy is inhabited by a large number of households. Throughout the book I use the terms *households, individuals,* and *agents* interchangeably. The Solow model makes relatively few assumptions about households, because their optimization problem is not explicitly modeled. This lack of optimization on the household side is the main difference between the Solow and the *neoclassical growth* models. The latter is the Solow model plus dynamic consumer (household) optimization. To fix ideas, you may want to assume that all households are identical, so that the economy trivially admits *a representative household*—meaning that the demand and labor supply side of the economy can be represented as if it resulted from the behavior of a single household. The representative household assumption is discussed in detail in Chapter 5.

What do we need to know about households in this economy? The answer is: not much. We have not yet endowed households with preferences (utility functions). Instead, for now, households are assumed to save a constant exogenous fraction $s \in (0, 1)$ of their disposable income—regardless of what else is happening in the economy. This assumption is the same as that used in basic Keynesian models and the Harrod-Domar model mentioned above. It is also at odds with reality. Individuals do not save a constant fraction of their incomes; if they did, then an announcement by the government that there will be a large tax increase next year should have no effect on their savings decisions, which seems both unreasonable and empirically incorrect. Nevertheless, the exogenous constant saving rate is a convenient starting point, and we will spend a lot of time in the rest of the book analyzing how consumers behave and make intertemporal choices.

The other key agents in the economy are firms. Firms, like consumers, are highly heterogeneous in practice. Even within a narrowly defined sector of an economy, no two firms are identical. But again for simplicity, let us start with an assumption similar to the representative household assumption, but now applied to firms: suppose that all firms in this economy have access to the same production function for the final good, or that the economy admits a *representative firm,* with a representative (or aggregate) production function. The conditions under which this represenitive firm assumption is reasonable are also discussed in Chapter 5. The aggregate production function for the unique final good is written as

$$Y(t) = F(K(t), L(t), A(t)), \tag{2.1}$$

where $Y(t)$ is the total amount of production of the final good at time t, $K(t)$ is the capital stock, $L(t)$ is total employment, and $A(t)$ is technology at time t. Employment can be measured in different ways. For example, we may want to think of $L(t)$ as corresponding to hours of employment or to number of employees. The capital stock $K(t)$ corresponds to the quantity of "machines" (or more specifically, equipment and structures) used in production, and it is typically measured in terms of the value of the machines. There are also multiple ways of thinking of capital (and equally many ways of specifying how capital comes into existence). Since the objective here is to start with a simple workable model, I make the rather sharp simplifying assumption that capital is the same as the final good of the economy. However, instead of being consumed, capital is used in the production process of more goods. To take a concrete example, think of the final good as "corn." Corn can be used both for consumption and as an input, as seed, for the production of more corn tomorrow. Capital then corresponds to the amount of corn used as seed for further production.

Technology, on the other hand, has no natural unit, and $A(t)$ is simply a *shifter* of the production function (2.1). For mathematical convenience, I often represent $A(t)$ in terms of a number, but it is useful to bear in mind that, at the end of the day, it is a representation of a more abstract concept. As noted in Chapter 1, we may often want to think of a broad notion of technology, incorporating the effects of the organization of production and of markets on the efficiency with which the factors of production are utilized. In the current model, $A(t)$ represents all these effects.

A major assumption of the Solow growth model (and of the neoclassical growth model we will study in Chapter 8) is that technology is *free:* it is publicly available as a nonexcludable, nonrival good. Recall that a good is *nonrival* if its consumption or use by others does not preclude an individual's consumption or use. It is *nonexcludable*, if it is impossible to prevent another person from using or consuming it. Technology is a good candidate for a nonexcludable, nonrival good; once the society has some knowledge useful for increasing the efficiency of production, this knowledge can be used by any firm without impinging on the use of it by others. Moreover, it is typically difficult to prevent firms from using this knowledge (at least once it is in the public domain and is not protected by patents). For example, once the society knows how to make wheels, everybody can use that knowledge to make wheels without diminishing the ability of others to do the same (thus making the knowledge to produce wheels nonrival). Moreover, unless somebody has a well-enforced patent on wheels, anybody can decide to produce wheels (thus making the knowhow to produce wheels nonexcludable). The implication of the assumptions that technology is nonrival and nonexcludable is that $A(t)$ is freely available to all potential firms in the economy and firms do not have to pay for making use of this technology. Departing from models in which technology is freely available is a major step toward understanding technological progress and will be our focus in Part IV.

As an aside, note that some authors use x_t or K_t when working with discrete time and reserve the notation $x(t)$ or $K(t)$ for continuous time. Since I go back and forth between continuous and discrete time, I use the latter notation throughout. When there is no risk of confusion, I drop the time arguments, but whenever there is the slightest risk of confusion, I err on the side of caution and include the time arguments.

Let us next impose the following standard assumptions on the aggregate production function.

Assumption 1 (Continuity, Differentiability, Positive and Diminishing Marginal Products, and Constant Returns to Scale) *The production function $F : \mathbb{R}^3_+ \to \mathbb{R}_+$ is twice differentiable in K and L, and satisfies*

$$F_K(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial K} > 0, \quad F_L(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial L} > 0,$$

$$F_{KK}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial K^2} < 0, \quad F_{LL}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial L^2} < 0.$$

Moreover, F exhibits constant returns to scale in K and L.

All of the components of Assumption 1 are important. First, the notation $F : \mathbb{R}^3_+ \to \mathbb{R}_+$ implies that the production function takes nonnegative arguments (i.e., $K, L \in \mathbb{R}_+$) and maps to nonnegative levels of output ($Y \in \mathbb{R}_+$). It is natural that the level of capital and the level of employment should be positive. Since A has no natural units, it could have been negative. But there is no loss of generality in restricting it to be positive. The second important aspect of Assumption 1 is that F is a continuous function in its arguments and is also differentiable. There are many interesting production functions that are not differentiable, and some interesting ones that are not even continuous. But working with differentiable functions makes it possible to use differential calculus, and the loss of some generality is a small price to pay for this convenience. Assumption 1 also specifies that marginal products are positive (so that the level of production increases with the amount of inputs); this restriction also rules out some potential production functions and can be relaxed without much complication (see Exercise 2.8). More importantly, Assumption 1 requires that the marginal products of both capital and labor are diminishing, that is, $F_{KK} < 0$ and $F_{LL} < 0$, so that more capital, holding everything else constant, increases output by less and less. And the same applies to labor. This property is sometimes also referred to as "diminishing returns" to capital and labor. The degree of diminishing returns to capital plays a very important role in many results of the basic growth model. In fact, the presence of diminishing returns to capital distinguishes the Solow growth model from its antecedent, the Harrod-Domar model (see Exercise 2.23).

The other important assumption is that of constant returns to scale. Recall that F exhibits *constant returns to scale* in K and L if it is *linearly homogeneous* (homogeneous of degree 1) in these two variables. More specifically:

Definition 2.1 *Let $K \in \mathbb{N}$. The function $g : \mathbb{R}^{K+2} \to \mathbb{R}$ is* homogeneous of degree m *in $x \in \mathbb{R}$ and $y \in \mathbb{R}$ if*

$$g(\lambda x, \lambda y, z) = \lambda^m g(x, y, z) \text{ for all } \lambda \in \mathbb{R}_+ \text{ and } z \in \mathbb{R}^K.$$

It can be easily verified that linear homogeneity implies that the production function F is concave, though not strictly so (see Exercise 2.2). Linearly homogeneous (constant returns to scale) production functions are particularly useful because of the following theorem.

Theorem 2.1 (Euler's Theorem) *Suppose that $g : \mathbb{R}^{K+2} \to \mathbb{R}$ is differentiable in $x \in \mathbb{R}$ and $y \in \mathbb{R}$, with partial derivatives denoted by g_x and g_y, and is homogeneous of degree m in x and y. Then*

$$mg(x, y, z) = g_x(x, y, z)x + g_y(x, y, z)y \text{ for all } x \in \mathbb{R}, y \in \mathbb{R}, \text{ and } z \in \mathbb{R}^K.$$

Moreover, $g_x(x, y, z)$ and $g_y(x, y, z)$ are themselves homogeneous of degree $m - 1$ in x and y.

Proof. We have that g is differentiable and

$$\lambda^m g(x, y, z) = g(\lambda x, \lambda y, z). \tag{2.2}$$

Differentiate both sides of (2.2) with respect to λ, which gives

$$m\lambda^{m-1} g(x, y, z) = g_x(\lambda x, \lambda y, z)x + g_y(\lambda x, \lambda y, z)y$$

for any λ. Setting $\lambda = 1$ yields the first result. To obtain the second result, differentiate both sides of (2.2) with respect to x:

$$\lambda g_x(\lambda x, \lambda y, z) = \lambda^m g_x(x, y, z).$$

Dividing both sides by λ establishes the desired result. ■

2.1.2 Endowments, Market Structure, and Market Clearing

The previous subsection has specified household behavior and the technology of production. The next step is to specify endowments, that is, the amounts of labor and capital that the economy starts with and who owns these endowments. We will then be in a position to investigate the allocation of resources in this economy. Resources (for a given set of households and production technology) can be allocated in many different ways, depending on the *institutional structure* of the society. Chapters 5–8 discuss how a social planner wishing to maximize a weighted average of the utilities of households might allocate resources, while Part VIII focuses on the allocation of resources favoring individuals who are politically powerful. The more familiar benchmark for the allocation of resources is to assume a specific set of market institutions, in particular, competitive markets. In competitive markets, households and firms act in a price-taking manner and pursue their own objectives, and prices clear markets. Competitive markets are a natural benchmark, and I start by assuming that all goods and factor markets are competitive. This is yet another assumption that is not totally innocuous. For example, both labor and capital markets have imperfections, with certain important implications for economic growth, and monopoly power in product markets plays a major role in Part IV. But these implications can be best appreciated by starting out with the competitive benchmark.

Before investigating trading in competitive markets, let us also specify the ownership of the endowments. Since competitive markets make sense only in the context of an economy with (at least partial) private ownership of assets and the means of production, it is natural to suppose that factors of production are owned by households. In particular, let us suppose that households own all labor, which they supply inelastically. Inelastic supply means that there is some endowment of labor in the economy, for example, equal to the population, $\bar{L}(t)$, and all of it will be supplied regardless of its (rental) price—as long as this price is nonnegative. The labor market clearing condition can then be expressed as:

$$L(t) = \bar{L}(t) \tag{2.3}$$

for all t, where $L(t)$ denotes the demand for labor (and also the level of employment). More generally, this equation should be written in complementary slackness form. In particular, let the rental price of labor or the wage rate at time t be $w(t)$, then the labor market clearing condition takes the form

$$L(t) \le \bar{L}(t),\ w(t) \ge 0 \quad \text{and} \quad \left(L(t) - \bar{L}(t)\right) w(t) = 0. \tag{2.4}$$

The complementary slackness formulation ensures that labor market clearing does not happen at a negative wage—or that if labor demand happens to be low enough, employment could be below $\bar{L}(t)$ at zero wage. However, this will not be an issue in most of the models studied in this book, because Assumption 1 and competitive labor markets ensure that wages are strictly positive (see Exercise 2.1). In view of this result, I use the simpler condition (2.3) throughout and denote both labor supply and employment at time t by $L(t)$.

The households also own the capital stock of the economy and rent it to firms. Let us denote the rental price of capital at time t by $R(t)$. The capital market clearing condition is similar to (2.3) and requires the demand for capital by firms to be equal to the supply of capital by households:

$$K(t) = \bar{K}(t),$$

where $\bar{K}(t)$ is the supply of capital by households and $K(t)$ is the demand by firms. Capital market clearing is straightforward to ensure in the class of models analyzed in this book. In particular, it is sufficient that the amount of capital $K(t)$ used in production at time t (from firms' optimization behavior) be consistent with households' endowments and saving behavior.

Let us take households' initial holdings of capital, $K(0) \geq 0$, as given (as part of the description of the environment). For now how this initial capital stock is distributed among the households is not important, since households' optimization decisions are not modeled explicitly and the economy is simply assumed to save a fraction s of its income. When we turn to models with household optimization below, an important part of the description of the environment will be to specify the preferences and the budget constraints of households.

At this point, I could also introduce the price of the final good at time t, say $P(t)$. But there is no need, since there is a choice of a numeraire commodity in this economy, whose price will be normalized to 1. In particular, as discussed in greater detail in Chapter 5, Walras's Law implies that the price of one of the commodities, the numeraire, should be normalized to 1. In fact, throughout I do something stronger and normalize the price of the final good to 1 in all periods. Ordinarily, one cannot choose more than one numeraire—otherwise, one would be fixing the relative price between the numeraires. But as explained in Chapter 5, we can build on an insight by Kenneth Arrow (Arrow, 1964) that it is sufficient to price *securities* (assets) that transfer one unit of consumption from one date (or state of the world) to another. In the context of dynamic economies, this implies that we need to keep track of an *interest rate* across periods, denoted by $r(t)$, which determines intertemporal prices and enables us to normalize the price of the final good to 1 within each period. Naturally we also need to keep track of the wage rate $w(t)$, which determines the price of labor relative to the final good at any date t.

This discussion highlights a central fact: all of the models in this book should be thought of as general equilibrium economies, in which different commodities correspond to the same good at different dates. Recall from basic general equilibrium theory that the same good at different dates (or in different states or localities) is a different commodity. Therefore, in almost all of the models in this book, there will be an infinite number of commodities, since time runs to infinity. This raises a number of special issues, which are discussed in Chapter 5 and later.

Returning to the basic Solow model, the next assumption is that capital depreciates, meaning that machines that are used in production lose some of their value because of wear and tear. In terms of the corn example above, some of the corn that is used as seed is no longer available for consumption or for use as seed in the following period. Let us assume that this depreciation takes an exponential form, which is mathematically very tractable. Thus capital depreciates (exponentially) at the rate $\delta \in (0, 1)$, so that out of 1 unit of capital this period, only $1 - \delta$ is left for next period. Though depreciation here stands for the wear and tear of the machinery, it can also represent the replacement of old machines by new ones in more realistic models (see Chapter 14).

The loss of part of the capital stock affects the interest rate (rate of return on savings) faced by households. Given the assumption of exponential depreciation at the rate δ and the normalization of the price of the final good to 1, the interest rate faced by the households is $r(t) = R(t) - \delta$, where recall that $R(t)$ is the rental price of capital at time t. A unit of final good can be consumed now or used as capital and rented to firms. In the latter case, a household receives $R(t)$ units of good in the next period as the rental price for its savings, but loses δ units of its capital holdings, since δ fraction of capital depreciates over time. Thus the household has given up one unit of commodity dated $t-1$ and receives $1 + r(t) = R(t) + 1 - \delta$ units of commodity dated t, so that $r(t) = R(t) - \delta$. The relationship between $r(t)$ and $R(t)$ explains the similarity between the symbols for the interest rate and the rental rate of capital. The interest rate faced by households plays a central role in the dynamic optimization decisions of households below. In the Solow model, this interest rate does not directly affect the allocation of resources.

2.1.3 Firm Optimization and Equilibrium

We are now in a position to look at the optimization problem of firms and the competitive equilibrium of this economy. Throughout the book I assume that the objective of firms is to maximize profits. Given the assumption that there is an aggregate production function, it is sufficient to consider the problem of a representative firm. Throughout, unless otherwise stated, I also assume that capital markets are functioning, so firms can rent capital in spot markets. For a given technology level $A(t)$, and given factor prices $R(t)$ and $w(t)$, the profit maximization problem of the representative firm at time t can be represented by the following static problem:

$$\max_{K\geq 0, L\geq 0} F(K, L, A(t)) - R(t)K - w(t)L. \tag{2.5}$$

When there are irreversible investments or costs of adjustments, as discussed, for example, in Section 7.8, the maximization problem of firms becomes dynamic. But in the absence of these features, maximizing profits separately at each date t is equivalent to maximizing the net present discounted value of profits. This feature simplifies the analysis considerably.

A couple of additional features are worth noting:

1. The maximization problem is set up in terms of aggregate variables, which, given the representative firm, is without any loss of generality.
2. There is nothing multiplying the F term, since the price of the final good has been normalized to 1. Thus the first term in (2.5) is the revenues of the representative firm (or the revenues of all of the firms in the economy).
3. This way of writing the problem already imposes competitive factor markets, since the firm is taking as given the rental prices of labor and capital, $w(t)$ and $R(t)$ (which are in terms of the numeraire, the final good).
4. This problem is concave, since F is concave (see Exercise 2.2).

An important aspect is that, because F exhibits constant returns to scale (Assumption 1), the maximization problem (2.5) does not have a well-defined solution (see Exercise 2.3); either there does not exist any (K, L) that achieves the maximum value of this program (which is infinity), or $K = L = 0$, or multiple values of (K, L) will achieve the maximum value of this program (when this value happens to be 0). This problem is related to the fact that in a world with constant returns to scale, the size of each individual firm is not determinate (only aggregates are determined). The same problem arises here because (2.5) is written without imposing the condition that factor markets should clear. A competitive equilibrium requires that all firms (and thus the representative firm) maximize profits and factor markets clear. In particular, the demands for labor and capital must be equal to the supplies of these factors at all times (unless the prices of these factors are equal to zero, which is ruled out by Assumption 1). This observation implies that the representative firm should make zero profits, since otherwise it would wish to hire arbitrarily large amounts of capital and labor exceeding the supplies, which are fixed. It also implies that total demand for labor, L, must be equal to the available supply of labor, $L(t)$. Similarly, the total demand for capital, K, should equal the total supply, $K(t)$. If this were not the case and $L < L(t)$, then there would be an excess supply of labor and the wage would be equal to zero. But this is not consistent with firm maximization, since given Assumption 1, the representative firm would then wish to hire an arbitrarily large amount of labor, exceeding the supply. This argument, combined with the fact that F is differentiable (Assumption 1), implies that given the supplies of capital and labor at time t, $K(t)$ and $L(t)$, factor prices must satisfy the following familiar conditions equating factor prices to marginal products:[1]

$$w(t) = F_L(K(t), L(t), A(t)), \tag{2.6}$$

and

$$R(t) = F_K(K(t), L(t), A(t)). \tag{2.7}$$

Euler's Theorem (Theorem 2.1) then verifies that at the prices (2.6) and (2.7), firms (or the representative firm) make zero profits.

Proposition 2.1 *Suppose Assumption 1 holds. Then, in the equilibrium of the Solow growth model, firms make no profits, and in particular,*

$$Y(t) = w(t)L(t) + R(t)K(t).$$

Proof. This result follows immediately from Theorem 2.1 for the case of constant returns to scale ($m = 1$). ■

Since firms make no profits in equilibrium, the ownership of firms does not need to be specified. All we need to know is that firms are profit-maximizing entities.

In addition to these standard assumptions on the production function, the following boundary conditions, the *Inada conditions,* are often imposed in the analysis of economic growth and macroeconomic equilibria.

Assumption 2 (Inada Conditions) *F satisfies the* Inada conditions

$$\lim_{K\to 0} F_K(K, L, A) = \infty \quad \text{and} \quad \lim_{K\to \infty} F_K(K, L, A) = 0 \text{ for all } L > 0 \text{ and all } A,$$

$$\lim_{L\to 0} F_L(K, L, A) = \infty \quad \text{and} \quad \lim_{L\to \infty} F_L(K, L, A) = 0 \text{ for all } K > 0 \text{ and all } A.$$

Moreover, $F(0, L, A) = 0$ for all L and A.

The role of these conditions—especially in ensuring the existence of *interior equilibria*—will become clear later in this chapter. They imply that the first units of capital and labor

1. An alternative way to derive (2.6) and (2.7) is to consider the cost minimization problem of the representative firm, which takes the form of minimizing $rK + wL$ with respect to K and L, subject to the constraint that $F(K, L, A) = Y$ for some level of output Y. This problem has a unique solution for any given level of Y. Then imposing market clearing, that is, $Y = F(K, L, A)$ with K and L corresponding to the supplies of capital and labor, yields (2.6) and (2.7).

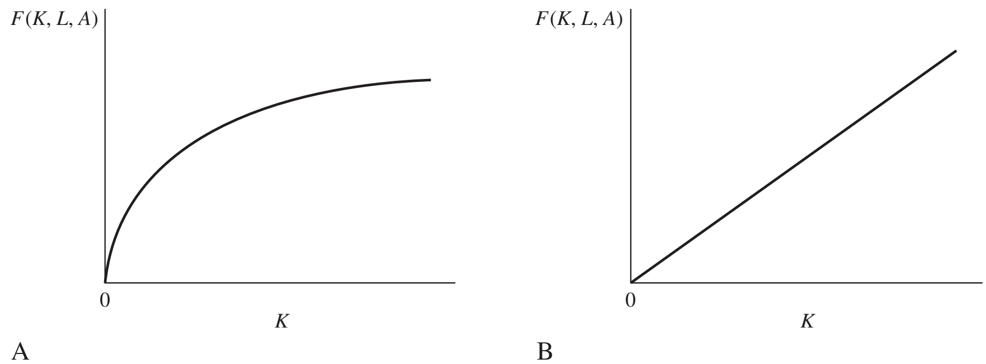

FIGURE 2.1 Production functions. (A) satisfies the Inada conditions in Assumption 2, while (B) does not.

are highly productive and that when capital or labor are sufficiently abundant, their marginal products are close to zero. The condition that $F(0, L, A) = 0$ for all L and A makes capital an essential input. This aspect of the assumption can be relaxed without any major implications for the results in this book. Figure 2.1 shows the production function $F(K, L, A)$ as a function of K, for given L and A, in two different cases; in panel A the Inada conditions are satisfied, while in panel B they are not.

I refer to Assumptions 1 and 2, which can be thought of as the neoclassical technology assumptions, throughout much of the book. For this reason, they are numbered independently from the equations, theorems, and proposition in this chapter.

2.2 The Solow Model in Discrete Time

I next present the dynamics of economic growth in the discrete-time Solow model.

2.2.1 Fundamental Law of Motion of the Solow Model

Recall that K depreciates exponentially at the rate δ, so that the law of motion of the capital stock is given by

$$K(t+1) = (1-\delta)\,K(t) + I(t), \tag{2.8}$$

where $I(t)$ is investment at time t.

From national income accounting for a closed economy, the total amount of final good in the economy must be either consumed or invested, thus

$$Y(t) = C(t) + I(t), \tag{2.9}$$

where $C(t)$ is consumption.[2] Using (2.1), (2.8), and (2.9), any feasible dynamic allocation in this economy must satisfy

$$K(t+1) \le F(K(t), L(t), A(t)) + (1-\delta)K(t) - C(t)$$

2. In addition, we can introduce government spending $G(t)$ on the right-hand side of (2.9). Government spending does not play a major role in the Solow growth model, thus its introduction is relegated to Exercise 2.7.

for $t = 0, 1, \ldots$. The question is to determine the equilibrium dynamic allocation among the set of feasible dynamic allocations. Here the behavioral rule that households save a constant fraction of their income simplifies the structure of equilibrium considerably (this is a behavioral rule, since it is not derived from the maximization of a well-defined utility function). One implication of this assumption is that any welfare comparisons based on the Solow model have to be taken with a grain of salt, since we do not know what the preferences of the households are.

Since the economy is closed (and there is no government spending), aggregate investment is equal to savings:

$$S(t) = I(t) = Y(t) - C(t).$$

The assumption that households save a constant fraction $s \in (0, 1)$ of their income can be expressed as

$$S(t) = sY(t), \tag{2.10}$$

which, in turn, implies that they consume the remaining $1 - s$ fraction of their income, and thus

$$C(t) = (1 - s)\, Y(t). \tag{2.11}$$

In terms of capital market clearing, (2.10) implies that the supply of capital for time $t + 1$ resulting from households' behavior can be expressed as $K(t + 1) = (1 - \delta)K(t) + S(t) = (1 - \delta)K(t) + sY(t)$. Setting supply and demand equal to each other and using (2.1) and (2.8) yields *the fundamental law of motion* of the Solow growth model:

$$K(t + 1) = sF(K(t), L(t), A(t)) + (1 - \delta)K(t). \tag{2.12}$$

This is a nonlinear difference equation. The equilibrium of the Solow growth model is described by (2.12) together with laws of motion for $L(t)$ and $A(t)$.

2.2.2 Definition of Equilibrium

The Solow model is a mixture of an old-style Keynesian model and a modern dynamic macroeconomic model. Households do not optimize when it comes to their savings or consumption decisions. Instead, their behavior is captured by (2.10) and (2.11). Nevertheless, firms still maximize profits, and factor markets clear. Thus it is useful to start defining equilibria in the way that is customary in modern dynamic macro models.

Definition 2.2 *In the basic Solow model for a given sequence of* $\{L(t), A(t)\}_{t=0}^{\infty}$ *and an initial capital stock* $K(0)$*, an* equilibrium path *is a sequence of capital stocks, output levels, consumption levels, wages, and rental rates* $\{K(t), Y(t), C(t), w(t), R(t)\}_{t=0}^{\infty}$ *such that* $K(t)$ *satisfies (2.12),* $Y(t)$ *is given by (2.1),* $C(t)$ *is given by (2.11), and* $w(t)$ *and* $R(t)$ *are given by (2.6) and (2.7), respectively.*

The most important point to note about Definition 2.2 is that an equilibrium is defined as an entire path of allocations and prices. An economic equilibrium does *not* refer to a static object; it specifies the entire path of behavior of the economy. Note also that Definition 2.2 incorporates the market clearing conditions, (2.6) and (2.7), into the definition of equilibrium. This practice is standard in macro and growth models. The alternative, which involves describing the equilibrium in more abstract terms, is discussed in Chapter 8 in the context of the neoclassical growth model (see, in particular, Definition 8.1).

2.2.3 Equilibrium without Population Growth and Technological Progress

It is useful to start with the following assumptions, which are relaxed later in this chapter:

1. There is no population growth; total population is constant at some level $L > 0$. Moreover, since households supply labor inelastically, this implies $L(t) = L$.
2. There is no technological progress, so that $A(t) = A$.

Let us define the capital-labor ratio of the economy as

$$k(t) \equiv \frac{K(t)}{L}, \tag{2.13}$$

which is a key object for the analysis. Now using the assumption of constant returns to scale, output (income) per capita, $y(t) \equiv Y(t)/L$, can be expressed as

$$\begin{aligned} y(t) &= F\left(\frac{K(t)}{L}, 1, A\right) \\ &\equiv f(k(t)). \end{aligned} \tag{2.14}$$

In other words, with constant returns to scale, output per capita is simply a function of the capital-labor ratio. Note that $f(k)$ here depends on A, so I could have written $f(k, A)$. I do not do this to simplify the notation and also because until Section 2.7, there will be no technological progress. Thus for now A is constant and can be normalized to $A = 1$.[3] The marginal product and the rental price of capital are then given by the derivative of F with respect to its first argument, which is $f'(k)$. The marginal product of labor and the wage rate are then obtained from Theorem 2.1, so that

$$\begin{aligned} R(t) &= f'(k(t)) > 0 \quad \text{and} \\ w(t) &= f(k(t)) - k(t)f'(k(t)) > 0. \end{aligned} \tag{2.15}$$

The fact that both factor prices are positive follows from Assumption 1, which ensures that the first derivatives of F with respect to capital and labor are always positive.

Example 2.1 (The Cobb-Douglas Production Function) *Let us consider the most common example of production function used in macroeconomics, the Cobb-Douglas production function. I hasten to add the caveat that even though the Cobb-Douglas form is convenient and widely used, it is also very special, and many interesting phenomena discussed later in this book are ruled out by this production function. The Cobb-Douglas production function can be written as*

$$\begin{aligned} Y(t) &= F(K(t), L(t), A(t)) \\ &= AK(t)^{\alpha}L(t)^{1-\alpha}, 0 < \alpha < 1. \end{aligned} \tag{2.16}$$

3. Later, when technological change is taken to be labor-augmenting, the term A can also be taken out, and the per capita production function can be written as $y = Af(k)$, with a slightly different definition of k as effective capital-labor ratio (see, e.g., (2.50) in Section 2.7).

It can easily be verified that this production function satisfies Assumptions 1 and 2, including the constant returns to scale feature imposed in Assumption 1. Dividing both sides by $L(t)$, the per capita production function in (2.14) becomes:

$$y(t) = Ak(t)^{\alpha},$$

where $y(t)$ again denotes output per worker and $k(t)$ is capital-labor ratio as defined in (2.13). The representation of factor prices as in (2.15) can also be verified. From the per capita production function representation, in particular (2.15), the rental price of capital can be expressed as

$$\begin{aligned} R(t) &= \frac{\partial Ak(t)^{\alpha}}{\partial k(t)}, \\ &= \alpha Ak(t)^{-(1-\alpha)}. \end{aligned}$$

Alternatively, in terms of the original production function (2.16), the rental price of capital in (2.7) is given by

$$\begin{aligned} R(t) &= \alpha AK(t)^{\alpha-1}L(t)^{1-\alpha} \\ &= \alpha Ak(t)^{-(1-\alpha)}, \end{aligned}$$

which is equal to the previous expression and thus verifies the form of the marginal product given in (2.15). Similarly, from (2.15),

$$\begin{aligned} w(t) &= Ak(t)^{\alpha} - \alpha Ak(t)^{-(1-\alpha)} \times k(t) \\ &= (1-\alpha)AK(t)^{\alpha}L(t)^{-\alpha}, \end{aligned}$$

which verifies the alternative expression for the wage rate in (2.6).

Returning to the analysis with the general production function, the per capita representation of the aggregate production function enables us to divide both sides of (2.12) by L to obtain the following simple difference equation for the evolution of the capital-labor ratio:

$$k(t+1) = sf(k(t)) + (1-\delta)k(t). \tag{2.17}$$

Since this difference equation is derived from (2.12), it also can be referred to as the *equilibrium difference equation* of the Solow model: it describes the equilibrium behavior of the key object of the model, the capital-labor ratio. The other equilibrium quantities can all be obtained from the capital-labor ratio $k(t)$.

At this point, let us also define a *steady-state equilibrium* for this model.

Definition 2.3 *A* steady-state equilibrium *without technological progress and population growth is an equilibrium path in which $k(t) = k^*$ for all t.*

In a steady-state equilibrium the capital-labor ratio remains constant. Since there is no population growth, this implies that the level of the capital stock will also remain constant. Mathematically, a steady-state equilibrium corresponds to a stationary point of the equilibrium difference equation (2.17). Most of the models in this book admit a steady-state equilibrium. This is also the case for this simple model.

The existence of a steady state can be seen by plotting the difference equation that governs the equilibrium behavior of this economy, (2.17), which is done in Figure 2.2. The thick curve

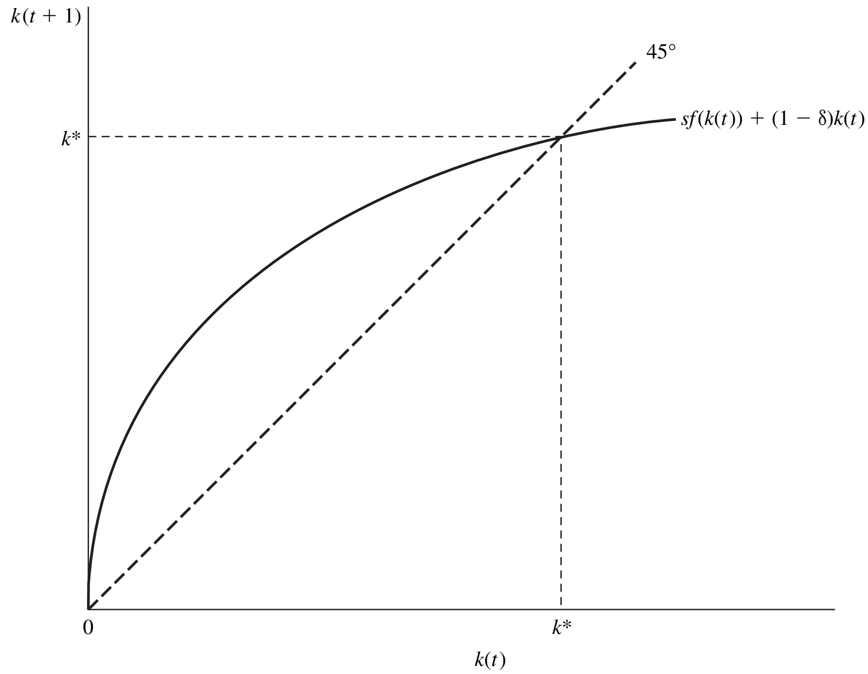

FIGURE 2.2 Determination of the steady-state capital-labor ratio in the Solow model without population growth and technological change.

represents the right-hand side of (2.17) and the dashed line corresponds to the 45° line. Their (positive) intersection gives the steady-state value of the capital-labor ratio k^*, which satisfies

$$\frac{f(k^*)}{k^*} = \frac{\delta}{s}. \tag{2.18}$$

Notice that in Figure 2.2 there is another intersection between (2.17) and the 45° line at $k = 0$. This second intersection occurs because, from Assumption 2, capital is an essential input, and thus $f(0) = 0$. Starting with $k(0) = 0$, there will then be no savings, and the economy will remain at $k = 0$. Nevertheless, I ignore this intersection throughout for a number of reasons. First, $k = 0$ is a steady-state equilibrium only when capital is an essential input and $f(0) = 0$. But as noted above, this assumption can be relaxed without any implications for the rest of the analysis, and when $f(0) > 0$, $k = 0$ is no longer a steady-state equilibrium. This is illustrated in Figure 2.3, which draws (2.17) for the case where $f(0) = \varepsilon$ for some $\varepsilon > 0$. Second, as we will see below, this intersection, even when it exists, is an unstable point; thus the economy would never travel toward this point starting with $K(0) > 0$ (or with $k(0) > 0$). Finally, and most importantly, this intersection holds no economic interest for us.[4]

An alternative visual representation shows the steady state as the intersection between a ray through the origin with slope δ (representing the function δk) and the function $sf(k)$. Figure 2.4, which illustrates this representation, is also useful for two other purposes. First, it depicts the levels of consumption and investment in a single figure. The vertical distance between the horizontal axis and the δk line at the steady-state equilibrium gives the amount of

4. Hakenes and Irmen (2006) show that even with $f(0) = 0$, the Inada conditions imply that in the continuous-time version of the Solow model $k = 0$ may not be the only equilibrium and the economy may move away from $k = 0$.

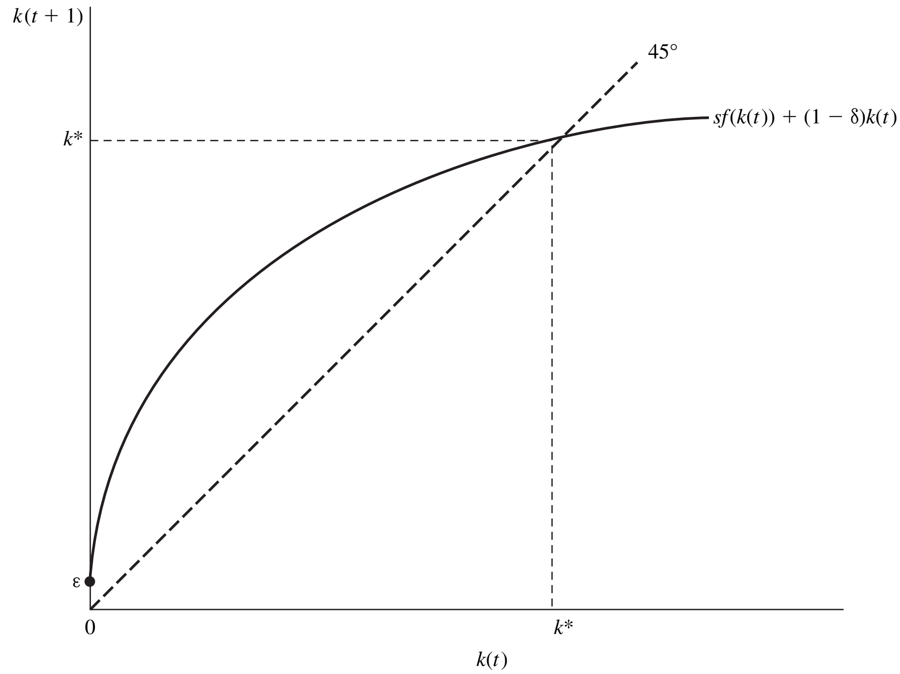

FIGURE 2.3 Unique steady state in the basic Solow model when $f(0) = \varepsilon > 0$.

investment per capita at the steady-state equilibrium (equal to δk^*), while the vertical distance between the function $f(k)$ and the δk line at k^* gives the level of consumption per capita. Clearly, the sum of these two terms make up $f(k^*)$. Second, Figure 2.4 also emphasizes that the steady-state equilibrium in the Solow model essentially sets investment, $sf(k)$, equal to the amount of capital that needs to be replenished, δk. This interpretation is particularly useful when population growth and technological change are incorporated.

This analysis therefore leads to the following proposition (with the convention that the intersection at $k = 0$ is being ignored even though $f(0) = 0$).

Proposition 2.2 *Consider the basic Solow growth model and suppose that Assumptions 1 and 2 hold. Then there exists a unique steady-state equilibrium where the capital-labor ratio* $k^* \in (0, \infty)$ *satisfies (2.18), per capita output is given by*

$$y^* = f(k^*), \tag{2.19}$$

and per capita consumption is given by

$$c^* = (1 - s)\, f(k^*). \tag{2.20}$$

Proof. The preceding argument establishes that any k^* that satisfies (2.18) is a steady state. To establish existence, note that from Assumption 2 (and from l'Hôpital's Rule, see Theorem A.21 in Appendix A), $\lim_{k\to 0} f(k)/k = \infty$ and $\lim_{k\to\infty} f(k)/k = 0$. Moreover, $f(k)/k$ is continuous from Assumption 1, so by the Intermediate Value Theorem (Theorem A.3) there exists k^* such that (2.18) is satisfied. To see uniqueness, differentiate $f(k)/k$ with respect to k, which gives

$$\frac{\partial(f(k)/k)}{\partial k} = \frac{f'(k)k - f(k)}{k^2} = -\frac{w}{k^2} < 0, \tag{2.21}$$

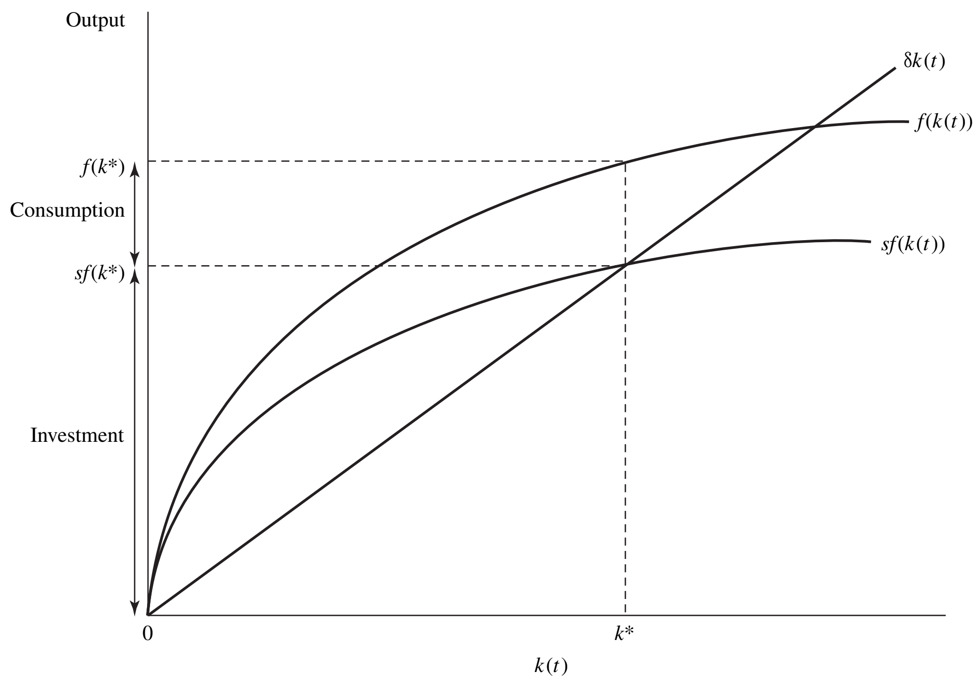

FIGURE 2.4 Investment and consumption in the steady-state equilibrium.

where the last equality in (2.21) uses (2.15). Since $f(k)/k$ is everywhere (strictly) decreasing, there can only exist a unique value k^* that satisfies (2.18). Equations (2.19) and (2.20) then follow by definition. ■

Through a series of examples, Figure 2.5 shows why Assumptions 1 and 2 cannot be dispensed with for establishing the existence and uniqueness results in Proposition 2.2. In the first two panels, the failure of Assumption 2 leads to a situation in which there is no steady-state equilibrium with positive activity, while in the third panel, the failure of Assumption 1 leads to nonuniqueness of steady states.

So far the model is very parsimonious: it does not have many parameters and abstracts from many features of the real world. An understanding of how cross-country differences in certain parameters translate into differences in growth rates or output levels is essential for our focus. This connection will be made in the next proposition. But before doing so, let us generalize the production function in one simple way and assume that

$$f(k) = A\tilde{f}(k),$$

where $A > 0$, so that A is a shift parameter, with greater values corresponding to greater productivity of factors. This type of productivity is referred to as "Hicks-neutral" (see below). For now, it is simply a convenient way of parameterizing productivity differences across countries. Since $f(k)$ satisfies the regularity conditions imposed above, so does $\tilde{f}(k)$.

Proposition 2.3 *Suppose Assumptions 1 and 2 hold and $f(k) = A\tilde{f}(k)$. Denote the steady-state level of the capital-labor ratio by $k^*(A, s, \delta)$ and the steady-state level of output by*

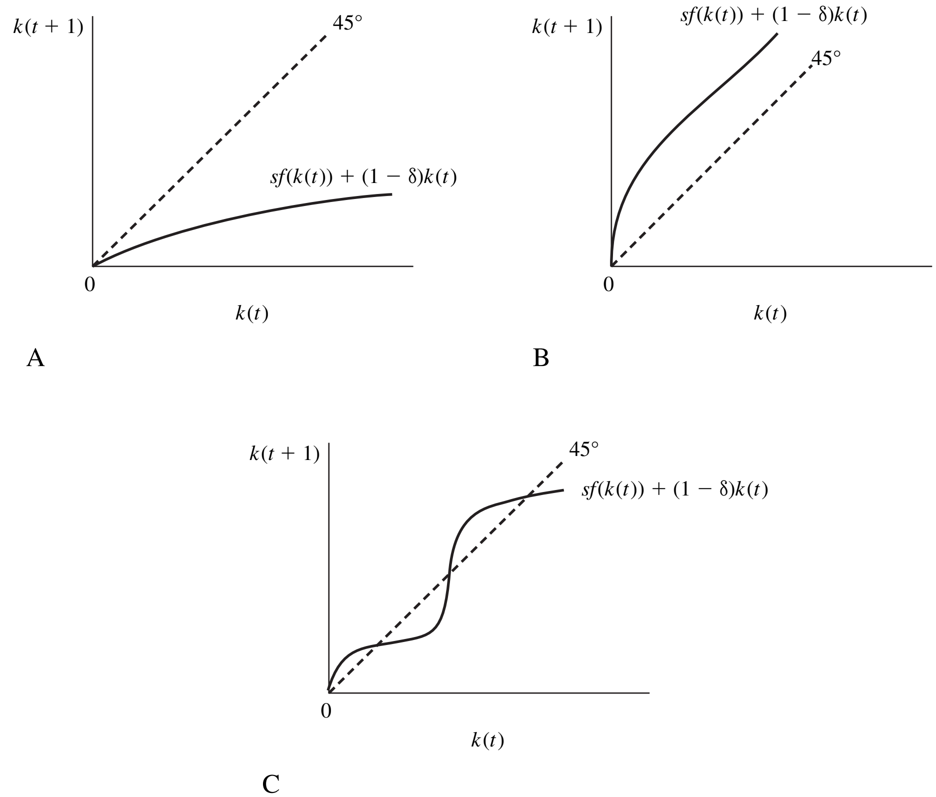

FIGURE 2.5 Examples of nonexistence and nonuniqueness of interior steady states when Assumptions 1 and 2 are not satisfied.

$y^*(A, s, \delta)$ when the underlying parameters are A, s, and δ. Then

$$\frac{\partial k^*(A, s, \delta)}{\partial A} > 0, \quad \frac{\partial k^*(A, s, \delta)}{\partial s} > 0, \quad \text{and } \frac{\partial k^*(A, s, \delta)}{\partial \delta} < 0;$$
$$\frac{\partial y^*(A, s, \delta)}{\partial A} > 0, \quad \frac{\partial y^*(A, s, \delta)}{\partial s} > 0, \quad \text{and } \frac{\partial y^*(A, s, \delta)}{\partial \delta} < 0.$$

Proof. The proof follows immediately by writing

$$\frac{\tilde{f}(k^*)}{k^*} = \frac{\delta}{As},$$

which holds for an open set of values of k^*, A, s, and δ. Now apply the Implicit Function Theorem (Theorem A.25) to obtain the results. For example,

$$\frac{\partial k^*}{\partial s} = \frac{\delta (k^*)^2}{s^2 w^*} > 0,$$

where $w^* = f(k^*) - k^* f'(k^*) > 0$. The other results follow similarly. ■

Therefore countries with higher saving rates and better technologies will have higher capital-labor ratios and will be richer. Those with greater (technological) depreciation will tend to have lower capital-labor ratios and will be poorer. All of the results in Proposition 2.3 are intuitive, and they provide us with a first glimpse of the potential determinants of the capital-labor ratios and output levels across countries.

The same comparative statics with respect to A and δ also apply to c^*. However, it is straightforward to see that c^* is not monotone in the saving rate (e.g., think of the extreme case where $s = 1$). In fact, there exists a unique saving rate, s_{gold}, referred to as the "golden rule" saving rate, which maximizes the steady-state level of consumption. Since we are treating the saving rate as an exogenous parameter and have not specified the objective function of households yet, we cannot say whether the golden rule saving rate is better than some other saving rate. It is nevertheless interesting to characterize what this golden rule saving rate corresponds to. To do this, let us first write the steady-state relationship between c^* and s and suppress the other parameters:

$$\begin{aligned} c^*(s) &= (1-s) f(k^*(s)) \\ &= f(k^*(s)) - \delta k^*(s), \end{aligned}$$

where the second equality exploits the fact that in steady state, $sf(k) = \delta k$. Now differentiating this second line with respect to s (again using the Implicit Function Theorem), we obtain

$$\frac{\partial c^*(s)}{\partial s} = [f'(k^*(s)) - \delta] \frac{\partial k^*}{\partial s}. \tag{2.22}$$

Let us define the golden rule saving rate s_{gold} to be such that $\partial c^*(s_{\text{gold}})/\partial s = 0$. The corresponding steady-state golden rule capital stock is defined as k^*_{gold}. These quantities and the relationship between consumption and the saving rate are plotted in Figure 2.6. The next proposition shows that s_{gold} and k^*_{gold} are uniquely defined.

Proposition 2.4 *In the basic Solow growth model, the highest level of steady-state consumption is reached for s_{gold}, with the corresponding steady-state capital level k^*_{gold} such that*

$$f'(k^*_{\text{gold}}) = \delta. \tag{2.23}$$

Proof. By definition $\partial c^*(s_{\text{gold}})/\partial s = 0$. From Proposition 2.3, $\partial k^*/\partial s > 0$; thus (2.22) can be equal to zero only when $f'(k^*(s_{\text{gold}})) = \delta$. Moreover, when $f'(k^*(s_{\text{gold}})) = \delta$, it can be verified that $\partial^2 c^*(s_{\text{gold}})/\partial s^2 < 0$, so $f'(k^*(s_{\text{gold}})) = \delta$ indeed corresponds to a local maximum. That $f'(k^*(s_{\text{gold}})) = \delta$ also yields the global maximum is a consequence of the following observations: for all $s \in [0, 1]$, we have $\partial k^*/\partial s > 0$, and moreover, when $s < s_{\text{gold}}$, $f'(k^*(s)) - \delta > 0$ by the concavity of f, so $\partial c^*(s)/\partial s > 0$ for all $s < s_{\text{gold}}$. By the converse argument, $\partial c^*(s)/\partial s < 0$ for all $s > s_{\text{gold}}$. Therefore only s_{gold} satisfies $f'(k^*(s)) = \delta$ and gives the unique global maximum of consumption per capita. ■

In other words, there exists a unique saving rate, s_{gold}, and also a unique corresponding capital-labor ratio, k^*_{gold}, given by (2.23), that maximize the level of steady-state consumption. When the economy is below k^*_{gold}, a higher saving rate will increase consumption, whereas when the economy is above k^*_{gold}, steady-state consumption can be raised by saving less. In the latter case, lower savings translate into higher consumption, because the capital-labor ratio of the economy is too high; households are investing too much and not consuming enough. This is the essence of the phenomenon of *dynamic inefficiency,* discussed in greater detail in Chapter 9. For now, recall that there is no explicit utility function here, so statements about inefficiency

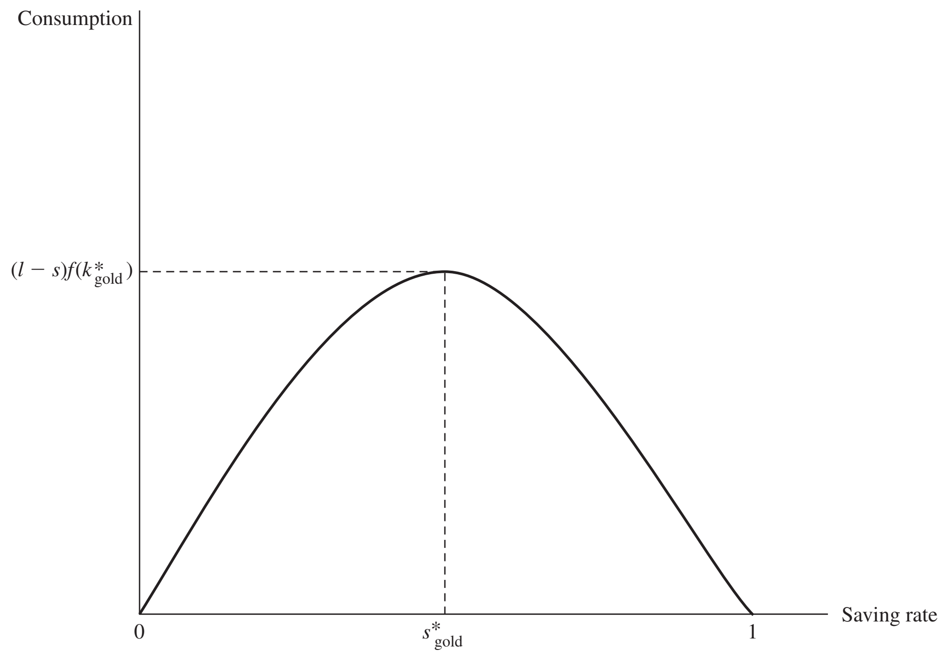

FIGURE 2.6 The golden rule level of saving rate, which maximizes steady-state consumption.

must be considered with caution. In fact, the reason this type of dynamic inefficiency does not generally apply when consumption-saving decisions are endogenized may already be apparent to many of you.

2.3 Transitional Dynamics in the Discrete-Time Solow Model

Proposition 2.2 establishes the existence of a unique steady-state equilibrium (with positive activity). Recall that an equilibrium path does not refer simply to the steady state but to the entire path of capital stock, output, consumption, and factor prices. This is an important point to bear in mind, especially since the term "equilibrium" is used differently in economics than in other disciplines. Typically, in engineering and the physical sciences, an equilibrium refers to a point of rest of a dynamical system, thus to what I have so far referred to as "the steady-state equilibrium." One may then be tempted to say that the system is in "disequilibrium" when it is away from the steady state. However, in economics, the non-steady-state behavior of an economy is also governed by market clearing and optimizing behavior of households and firms. Most economies spend much of their time in non-steady-state situations. Thus we are typically interested in the entire dynamic equilibrium path of the economy, not just in its steady state.

To determine what the equilibrium path of our simple economy looks like, we need to study the transitional dynamics of the equilibrium difference equation (2.17) starting from an arbitrary capital-labor ratio, $k(0) > 0$. Of special interest are the answers to the questions of whether the economy will tend to this steady state starting from an arbitrary capital-labor ratio and how it will behave along the transition path. Recall that the total amount of capital at the beginning of the economy, $K(0) > 0$, is taken as a state variable, while for now, the supply of labor L is fixed. Therefore at time $t = 0$, the economy starts with an arbitrary capital-labor ratio $k(0) = K(0)/L > 0$ as its initial value and then follows the law of motion given by the difference equation (2.17). Thus the question is whether (2.17) will take us to the unique steady state starting from an arbitrary initial capital-labor ratio.

Before answering this question, recall some definitions and key results from the theory of dynamical systems. Appendix B provides more details and a number of further results. Consider the nonlinear system of autonomous difference equations,

$$\mathbf{x}(t+1) = \mathbf{G}(\mathbf{x}(t)), \tag{2.24}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$ (where $n \in \mathbb{R}$). Let $\mathbf{x}^*$ be a *fixed point* of the mapping $\mathbf{G}(\cdot)$, that is,

$$\mathbf{x}^* = \mathbf{G}(\mathbf{x}^*).$$

I refer to $\mathbf{x}^*$ as a "steady state" of the difference equation (2.24).[5] The relevant notion of stability is introduced in the next definition.

Definition 2.4 *A steady state* $\mathbf{x}^*$ *is* locally asymptotically stable *if there exists an open set* $B(\mathbf{x}^*)$ *containing* $\mathbf{x}^*$ *such that for any solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$ *to (2.24) with* $\mathbf{x}(0) \in B(\mathbf{x}^*)$, $\mathbf{x}(t) \to \mathbf{x}^*$. *Moreover,* $\mathbf{x}^*$ *is* globally asymptotically stable *if for all* $\mathbf{x}(0) \in \mathbb{R}^n$, *for any solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$, $\mathbf{x}(t) \to \mathbf{x}^*$.

The next theorem provides the main results on the stability properties of systems of linear difference equations. The following theorems are special cases of the results presented in Appendix B.

Theorem 2.2 (Stability for Systems of Linear Difference Equations) *Consider the following linear difference equation system:*

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}, \tag{2.25}$$

with initial value $\mathbf{x}(0)$, *where* $\mathbf{x}(t) \in \mathbb{R}^n$ *for all* t, $\mathbf{A}$ *is an* $n \times n$ *matrix, and* $\mathbf{b}$ *is a* $n \times 1$ *column vector. Let* $\mathbf{x}^*$ *be the steady state of the difference equation given by* $\mathbf{A}\mathbf{x}^* + \mathbf{b} = \mathbf{x}^*$. *Suppose that all of the eigenvalues of* $\mathbf{A}$ *are strictly inside the unit circle in the complex plane. Then the steady state of the difference equation (2.25),* $\mathbf{x}^*$, *is globally (asymptotically) stable, in the sense that starting from any* $\mathbf{x}(0) \in \mathbb{R}^n$, *the unique solution* $\{\mathbf{x}(t)\}_{t=0}^{\infty}$ *satisfies* $\mathbf{x}(t) \to \mathbf{x}^*$.

Unfortunately, much less can be said about nonlinear systems, but the following is a standard local stability result.

Theorem 2.3 (Local Stability for Systems of Nonlinear Difference Equations) *Consider the following nonlinear autonomous system:*

$$\mathbf{x}(t+1) = \mathbf{G}(\mathbf{x}(t)), \tag{2.26}$$

with initial value $\mathbf{x}(0)$, *where* $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$. *Let* $\mathbf{x}^*$ *be a steady state of this system, that is,* $\mathbf{G}(\mathbf{x}^*) = \mathbf{x}^*$, *and suppose that* $\mathbf{G}$ *is differentiable at* $\mathbf{x}^*$. *Define*

$$\mathbf{A} \equiv D\mathbf{G}(\mathbf{x}^*),$$

5. Various other terms are used to describe $\mathbf{x}^*$, for example, "equilibrium point" or "critical point." Since these other terms have different meanings in economics, I refer to $\mathbf{x}^*$ as a steady state throughout.

where $D\mathbf{G}$ denotes the matrix of partial derivatives (Jacobian) of $\mathbf{G}$. Suppose that all of the eigenvalues of $\mathbf{A}$ are strictly inside the unit circle. Then the steady state of the difference equation (2.26), $\mathbf{x}^*$, is locally (asymptotically) stable, in the sense that there exists an open neighborhood of $\mathbf{x}^*$, $\mathbf{B}(\mathbf{x}^*) \subset \mathbb{R}^n$, such that starting from any $\mathbf{x}(0) \in \mathbf{B}(\mathbf{x}^*)$, $\mathbf{x}(t) \to \mathbf{x}^*$.

An immediate corollary of Theorem 2.3 is the following useful result.

Corollary 2.1

1. *Let $x(t), a, b \in \mathbb{R}$. If $|a| < 1$, then the unique steady state of the linear difference equation $x(t+1) = ax(t) + b$ is globally (asymptotically) stable, in the sense that $x(t) \to x^* = b/(1-a)$.*
2. *Let $g : \mathbb{R} \to \mathbb{R}$ be differentiable in the neighborhood of the steady state x^*, defined by $g(x^*) = x^*$, and suppose that $\left|g'(x^*)\right| < 1$. Then the steady state x^* of the nonlinear difference equation $x(t+1) = g(x(t))$ is locally (asymptotically) stable. Moreover, if g is continuously differentiable and satisfies $\left|g'(x)\right| < 1$ for all $x \in \mathbb{R}$, then x^* is globally (asymptotically) stable.*

Proof. The first part follows immediately from Theorem 2.2. The local stability of g in the second part follows from Theorem 2.3. Global stability follows since

$$\begin{aligned} |x(t+1) - x^*| &= |g(x(t)) - g(x^*)| \\ &= \left| \int_{x^*}^{x(t)} g'(x)dx \right| \\ &< |x(t) - x^*|, \end{aligned}$$

where the second line follows from the Fundamental Theorem of Calculus (Theorem B.2 in Appendix B), and the last inequality uses the hypothesis that $\left|g'(x)\right| < 1$ for all $x \in \mathbb{R}$. This implies that for any $x(0) < x^*$, $\{x(t)\}_{t=0}^{\infty}$ is an increasing sequence. Since $\left|g'(x)\right| < 1$, there cannot exist $x' \neq x^*$ such that $x' = g(x')$, and moreover $\{x(t)\}_{t=0}^{\infty}$ is bounded above by x^*. It therefore converges to x^*. The argument for the case where $x(0) > x^*$ is identical. ■

We can now apply Corollary 2.1 to the equilibrium difference equation (2.17) of the Solow model to establish the local stability of the steady-state equilibrium. Global stability does not directly follow form Corollary 2.1 (since the equivalent of $\left|g'(x)\right| < 1$ for all x is not true), but a slightly different argument can be used to prove this property.

Proposition 2.5 *Suppose that Assumptions 1 and 2 hold. Then the steady-state equilibrium of the Solow growth model described by the difference equation (2.17) is globally asymptotically stable, and starting from any $k(0) > 0$, $k(t)$ monotonically converges to k^*.*

Proof. Let $g(k) \equiv sf(k) + (1-\delta)k$. First observe that $g'(k)$ exists and is always strictly positive, that is, $g'(k) > 0$ for all k. Next, from (2.17),

$$k(t+1) = g(k(t)), \tag{2.27}$$

with a unique steady state at k^*. From (2.18), the steady-state capital k^* satisfies $\delta k^* = sf(k^*)$, or

$$k^* = g(k^*). \tag{2.28}$$

Now recall that $f(\cdot)$ is concave and differentiable from Assumption 1 and satisfies $f(0) = 0$ from Assumption 2. For any strictly concave differentiable function, we have (recall Fact A.23 in Appendix A):

$$f(k) > f(0) + kf'(k) = kf'(k). \tag{2.29}$$

Since (2.29) implies that $\delta = sf(k^*)/k^* > sf'(k^*)$, we have $g'(k^*) = sf'(k^*) + 1 - \delta < 1$. Therefore

$$g'(k^*) \in (0, 1).$$

Corollary 2.1 then establishes local asymptotic stability.

To prove global stability, note that for all $k(t) \in (0, k^*)$,

$$\begin{aligned} k(t+1) - k^* &= g(k(t)) - g(k^*) \\ &= -\int_{k(t)}^{k^*} g'(k)dk, \\ &< 0, \end{aligned}$$

where the first line follows by subtracting (2.28) from (2.27), the second line again uses the Fundamental Theorem of Calculus (Theorem B.2), and the third line follows from the observation that $g'(k) > 0$ for all k. Next, (2.17) also implies

$$\begin{aligned} \frac{k(t+1) - k(t)}{k(t)} &= s\frac{f(k(t))}{k(t)} - \delta \\ &> s\frac{f(k^*)}{k^*} - \delta \\ &= 0, \end{aligned}$$

where the second line uses the fact that $f(k)/k$ is decreasing in k (from (2.29)) and the last line uses the definition of k^*. These two arguments together establish that for all $k(t) \in (0, k^*)$, $k(t+1) \in (k(t), k^*)$. Therefore $\{k(t)\}_{t=0}^{\infty}$ is monotonically increasing and is bounded above by k^*. Moreover, since k^* is the unique steady state (with $k > 0$), there exists no $k' \in (0, k^*)$ such that $k(t+1) = k(t) = k'$ for any t. Therefore $\{k(t)\}_{t=0}^{\infty}$ must monotonically converge to k^*. An identical argument implies that for all $k(t) > k^*$, $k(t+1) \in (k^*, k(t))$ and establishes monotonic convergence starting from $k(0) > k^*$. This completes the proof of global stability. ■

This stability result can be seen diagrammatically in Figure 2.7. Starting from initial capital stock $k(0) > 0$, which is below the steady-state level k^*, the economy grows toward k^* and experiences *capital deepening*—meaning that the capital-labor ratio increases. Together with capital deepening comes growth of per capita income. If instead the economy were to start with $k'(0) > k^*$, it would reach the steady state by decumulating capital and contracting (i.e., by experiencing negative growth).

The following proposition is an immediate corollary of Proposition 2.5.

Proposition 2.6 *Suppose that Assumptions 1 and 2 hold, and $k(0) < k^*$. Then $\{w(t)\}_{t=0}^{\infty}$ is an increasing sequence, and $\{R(t)\}_{t=0}^{\infty}$ is a decreasing sequence. If $k(0) > k^*$, the opposite results apply.*

Proof. See Exercise 2.9. ■

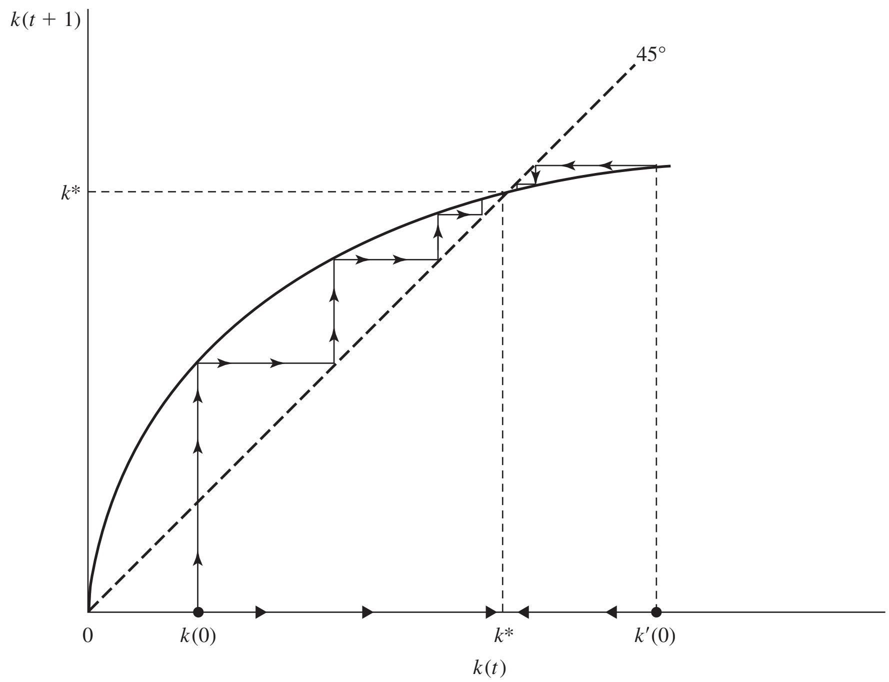

FIGURE 2.7 Transitional dynamics in the basic Solow model.

Recall that when the economy starts with too little capital relative to its labor supply, the capital-labor ratio will increase. Thus the marginal product of capital will fall due to diminishing returns to capital and the wage rate will increase. Conversely, if it starts with too much capital, it will decumulate capital, and in the process the wage rate will decline and the rate of return to capital will increase.

The analysis has established that the Solow growth model has a number of nice properties: unique steady state, global (asymptotic) stability, and finally, simple and intuitive comparative statics. Yet so far it has no growth. The steady state is the point at which there is no growth in the capital-labor ratio, no more capital deepening, and no growth in output per capita. Consequently, the basic Solow model (without technological progress) can only generate economic growth along the transition path to the steady state (starting with $k(0) < k^*$). However this growth is not sustained: it slows down over time and eventually comes to an end. Section 2.7 shows that the Solow model can incorporate economic growth by allowing exogenous technological change. Before doing this, it is useful to look at the relationship between the discrete- and continuous-time formulations.

2.4 The Solow Model in Continuous Time

2.4.1 From Difference to Differential Equations

Recall that the time periods $t = 0, 1, \ldots$ can refer to days, weeks, months, or years. In some sense, the time unit is not important. This arbitrariness suggests that perhaps it may be more convenient to look at dynamics by making the time unit as small as possible, that is, by going to continuous time. While much of modern macroeconomics (outside of growth theory) uses discrete-time models, many growth models are formulated in continuous time. The continuous-time setup has a number of advantages, since some pathological results of discrete-time models disappear when using continuous time (see Exercise 2.21). Moreover, continuous-time models have more flexibility in the analysis of dynamics and allow explicit-form solutions in a wider set of circumstances. These considerations motivate the detailed study of both the discrete- and continuous-time versions of the basic models in this book.

Let us start with a simple difference equation:

$$x(t+1) - x(t) = g(x(t)). \tag{2.30}$$

This equation states that between time t and $t+1$, the absolute growth in x is given by $g(x(t))$. Imagine that time is more finely divisible than that captured by our discrete indices, $t = 0, 1, \ldots$. In the limit, we can think of time as being as finely divisible as we would like, so that $t \in \mathbb{R}_+$. In that case, (2.30) gives us information about how the variable x changes between two discrete points in time, t and $t+1$. Between these time periods, we do not know how x evolves. However, if t and $t+1$ are not too far apart, the following approximation is reasonable:

$$x(t+\Delta t) - x(t) \simeq \Delta t \cdot g(x(t))$$

for any $\Delta t \in [0, 1]$. When $\Delta t = 0$, this equation is just an identity. When $\Delta t = 1$, it gives (2.30). In between it is a linear approximation. This approximation will be relatively accurate if the distance between t and $t+1$ is not very large, so that $g(x) \simeq g(x(t))$ for all $x \in [x(t), x(t+1)]$ (however, you should also convince yourself that this approximation could in fact be quite bad if the function g is highly nonlinear, in which case its behavior changes significantly between $x(t)$ and $x(t+1)$). Now divide both sides of this equation by Δt, and take limits to obtain

$$\lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t} = \dot{x}(t) \simeq g(x(t)), \tag{2.31}$$

where, as throughout the book, I use the dot notation to denote time derivatives, $\dot{x}(t) \equiv dx(t)/dt$. Equation (2.31) is a differential equation representing the same dynamics as the difference equation (2.30) for the case in which the distance between t and $t+1$ is small.

2.4.2 The Fundamental Equation of the Solow Model in Continuous Time

We can now repeat all of the analysis so far using the continuous-time representation. Nothing has changed on the production side, so we continue to have (2.6) and (2.7) as the factor prices, but now these refer to instantaneous rental rates. For example, $w(t)$ is the flow of wages that workers receive at instant t. Savings are again given by

$$S(t) = sY(t),$$

while consumption is still given by (2.11).

Let us also introduce population growth into this model and assume that the labor force $L(t)$ grows proportionally, that is,

$$L(t) = \exp(nt)L(0). \tag{2.32}$$

The purpose of doing so is that in many of the classical analyses of economic growth, population growth plays an important role, so it is useful to see how it affects the equilibrium here. There is still no technological progress.

Recall that

$$k(t) \equiv \frac{K(t)}{L(t)},$$

which implies that

$$\frac{\dot{k}(t)}{k(t)} = \frac{\dot{K}(t)}{K(t)} - \frac{\dot{L}(t)}{L(t)},$$
$$= \frac{\dot{K}(t)}{K(t)} - n,$$

where I used the fact that, from (2.32), $\dot{L}(t)/L(t) = n$. From the limiting argument leading to equation (2.31) in the previous subsection, the law of motion of the capital stock is given by

$$\dot{K}(t) = sF(K(t), L(t), A(t)) - \delta K(t).$$

Using the definition of $k(t)$ as the capital-labor ratio and the constant returns to scale properties of the production function, the fundamental law of motion of the Solow model in continuous time is obtained as

$$\frac{\dot{k}(t)}{k(t)} = s\frac{f(k(t))}{k(t)} - (n + \delta), \tag{2.33}$$

where, following usual practice, I have transformed the left-hand side to the proportional change in the capital-labor ratio by dividing both sides by $k(t)$.[6]

Definition 2.5 *In the basic Solow model in continuous time with population growth at the rate n, no technological progress and an initial capital stock $K(0)$, an* equilibrium path *is given by paths (sequences) of capital stocks, labor, output levels, consumption levels, wages, and rental rates $[K(t), L(t), Y(t), C(t), w(t), R(t)]_{t=0}^{\infty}$ such that $L(t)$ satisfies (2.32), $k(t) \equiv K(t)/L(t)$ satisfies (2.33), $Y(t)$ is given by (2.1), $C(t)$ is given by (2.11), and $w(t)$ and $R(t)$ are given by (2.6) and (2.7), respectively.*

As before, a steady-state equilibrium involves $k(t)$ remaining constant at some level k^*.

It is easy to verify that the equilibrium differential equation (2.33) has a unique steady state at k^*, which is given by a slight modification of (2.18) to incorporate population growth:

$$\frac{f(k^*)}{k^*} = \frac{n + \delta}{s}. \tag{2.34}$$

In other words, going from discrete to continuous time has not changed any of the basic economic features of the model. Thus the steady state can again be plotted in a diagram similar to Figure 2.1 except that it now also incorporates population growth. This is done in Figure 2.8, which also highlights that the logic of the steady state is the same with population growth as it was without population growth. The amount of investment, $sf(k)$, is used to replenish the capital-labor ratio, but now there are two reasons for replenishments. The capital stock depreciates exponentially at the flow rate δ. In addition, the capital stock must also increase as

6. Throughout I adopt the notation $[x(t)]_{t=0}^{\infty}$ to denote the continuous-time path of variable $x(t)$. An alternative notation often used in the literature is $(x(t); t \geq 0)$. I prefer the former both because it is slightly more compact and also because it is more similar to the discrete-time notation for the time path of a variable, $\{x(t)\}_{t=0}^{\infty}$. When referring to $[x(t)]_{t=0}^{\infty}$, I use the terms "path," "sequence," and "function (of time t)" interchangeably.

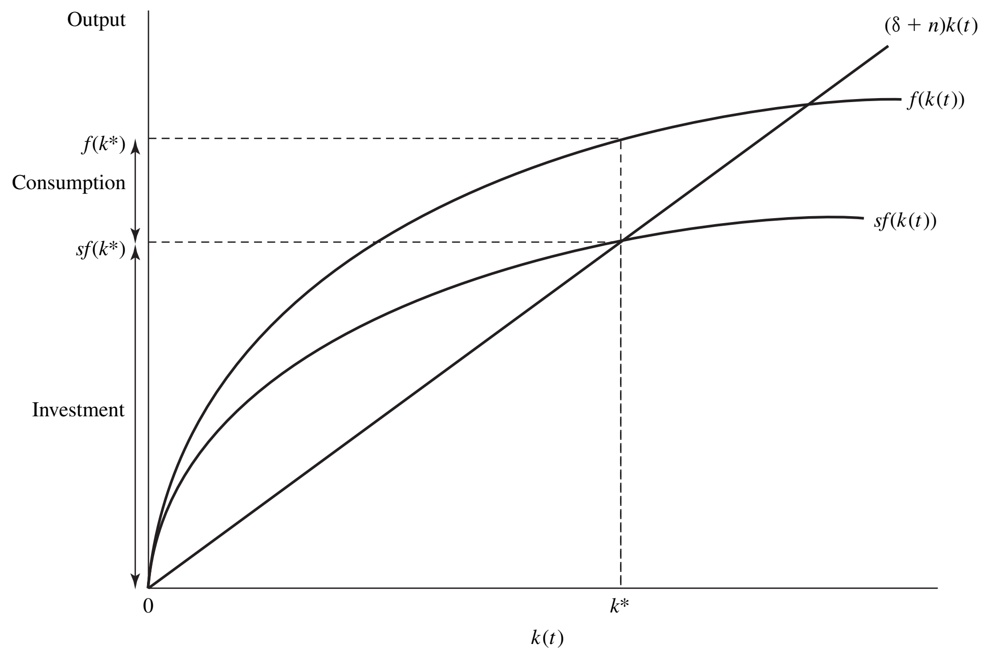

FIGURE 2.8 Investment and consumption in the steady-state equilibrium with population growth.

population grows to maintain the capital-labor ratio at a constant level. The amount of capital that needs to be replenished is therefore $(n + \delta)k$.

Proposition 2.7 *Consider the basic Solow growth model in continuous time and suppose that Assumptions 1 and 2 hold. Then there exists a unique steady-state equilibrium where the capital-labor ratio is equal to* $k^* \in (0, \infty)$ *and satisfies (2.34), per capita output is given by*

$$y^* = f(k^*),$$

and per capita consumption is given by

$$c^* = (1 - s)f(k^*).$$

Proof. See Exercise 2.5. ■

Moreover, again defining $f(k) = A\tilde{f}(k)$, the following proposition holds.

Proposition 2.8 *Suppose Assumptions 1 and 2 hold and* $f(k) = A\tilde{f}(k)$. *Denote the steady-state equilibrium level of the capital-labor ratio by* $k^*(A, s, \delta, n)$ *and the steady-state level of output by* $y^*(A, s, \delta, n)$ *when the underlying parameters are given by* A, s, δ, *and* n. *Then we have*

$$\frac{\partial k^*(A,s,\delta,n)}{\partial A} > 0, \quad \frac{\partial k^*(A,s,\delta,n)}{\partial s} > 0, \quad \frac{\partial k^*(A,s,\delta,n)}{\partial \delta} < 0, \quad \text{and} \quad \frac{\partial k^*(A,s,\delta,n)}{\partial n} < 0;$$
$$\frac{\partial y^*(A,s,\delta,n)}{\partial A} > 0, \quad \frac{\partial y^*(A,s,\delta,n)}{\partial s} > 0, \quad \frac{\partial y^*(A,s,\delta,n)}{\partial \delta} < 0, \quad \text{and} \quad \frac{\partial y^*(A,s,\delta,n)}{\partial n} < 0.$$

Proof. See Exercise 2.6. ■

The new result relative to the earlier comparative static proposition (Proposition 2.3) is that now a higher population growth rate, n, also reduces the capital-labor ratio and output per capita. The reason for this is simple: a higher population growth rate means there is more labor to use the existing amount of capital, which only accumulates slowly, and consequently the equilibrium capital-labor ratio ends up lower. This result implies that countries with higher population growth rates will have lower incomes per person (or per worker).

2.5 Transitional Dynamics in the Continuous-Time Solow Model

The analysis of transitional dynamics and stability with continuous time yields similar results to those in Section 2.3, but the analysis is slightly simpler. Let us first recall the basic results on the stability of systems of differential equations. Once again, further details are contained in Appendix B.

Theorem 2.4 (Stability of Linear Differential Equations) *Consider the following linear differential equation system:*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b} \tag{2.35}$$

with initial value $\mathbf{x}(0)$*, where* $\mathbf{x}(t) \in \mathbb{R}^n$ *for all* t*,* $\mathbf{A}$ *is an* $n \times n$ *matrix, and* $\mathbf{b}$ *is a* $n \times 1$ *column vector. Let* $\mathbf{x}^*$ *be the steady state of the system given by* $\mathbf{A}\mathbf{x}^* + \mathbf{b} = 0$*. Suppose that all eigenvalues of* $\mathbf{A}$ *have negative real parts. Then the steady state of the differential equation (2.35)* $\mathbf{x}^*$ *is globally asymptotically stable, in the sense that starting from any* $\mathbf{x}(0) \in \mathbb{R}^n$*,* $\mathbf{x}(t) \to \mathbf{x}^*$*.*

Theorem 2.5 (Local Stability of Nonlinear Differential Equations) *Consider the following nonlinear autonomous differential equation:*

$$\dot{\mathbf{x}}(t) = \mathbf{G}(\mathbf{x}(t)) \tag{2.36}$$

with initial value $\mathbf{x}(0)$*, where* $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$*. Let* $\mathbf{x}^*$ *be a steady state of this system, that is,* $\mathbf{G}(\mathbf{x}^*) = 0$*, and suppose that* $\mathbf{G}$ *is differentiable at* $\mathbf{x}^*$*. Define*

$$\mathbf{A} \equiv D\mathbf{G}(\mathbf{x}^*),$$

and suppose that all eigenvalues of $\mathbf{A}$ *have negative real parts. Then the steady state of the differential equation (2.36),* $\mathbf{x}^*$*, is locally asymptotically stable, in the sense that there exists an open neighborhood of* $\mathbf{x}^*$*,* $\mathbf{B}(\mathbf{x}^*) \subset \mathbb{R}^n$*, such that starting from any* $\mathbf{x}(0) \in \mathbf{B}(\mathbf{x}^*)$*,* $\mathbf{x}(t) \to \mathbf{x}^*$*.*

Once again an immediate corollary is as follows.

Corollary 2.2

1. *Let* $x(t) \in \mathbb{R}$*. Then the steady state of the linear differential equation* $\dot{x}(t) = ax(t)$ *is globally asymptotically stable (in the sense that* $x(t) \to 0$*) if* $a < 0$*.*
2. *Let* $g : \mathbb{R} \to \mathbb{R}$ *be differentiable in the neighborhood of the steady state* x^* *defined by* $g(x^*) = 0$ *and suppose that* $g'(x^*) < 0$*. Then the steady state of the nonlinear differential equation* $\dot{x}(t) = g(x(t))$*,* x^**, is locally asymptotically stable.*
3. *Let* $g : \mathbb{R} \to \mathbb{R}$ *be continuously differentiable. Suppose that* $g(x^*) = 0$ *and that* $g(x) < 0$ *for all* $x > x^*$*and* $g(x) > 0$ *for all* $x < x^*$*. Then the steady state of the nonlinear differential equation* $\dot{x}(t) = g(x(t))$*,* x^**, is globally asymptotically stable, that is, starting with any* $x(0)$*,* $x(t) \to x^*$*.*

Proof. See Exercise 2.10. ■

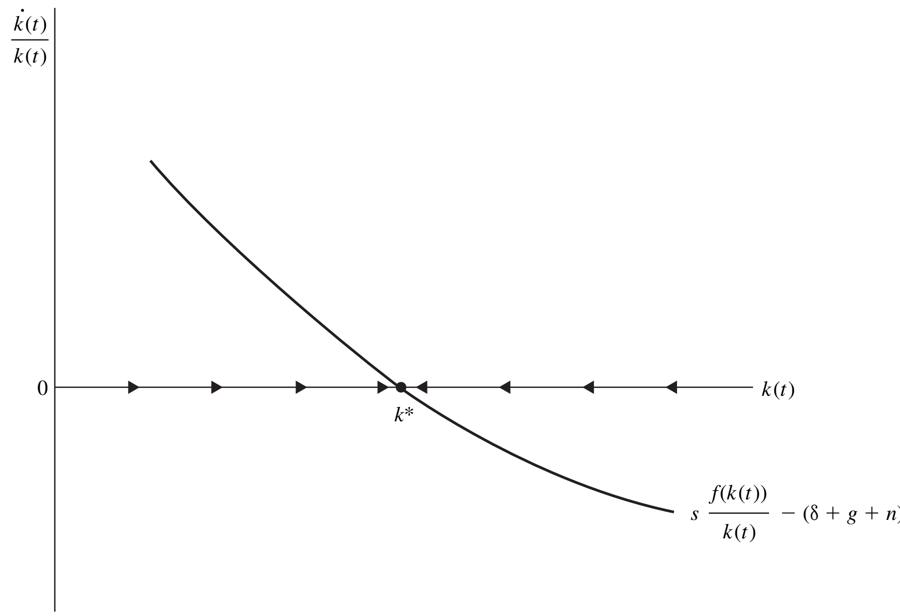

FIGURE 2.9 Dynamics of the capital-labor ratio in the basic Solow model.

Notice that the equivalent of part 3 of Corollary 2.2 is not true in discrete time. The implications of this observation are illustrated in Exercise 2.21.

In view of these results, Proposition 2.5 has a straightforward generalization of the results for discrete time.

Proposition 2.9 *Suppose that Assumptions 1 and 2 hold. Then the basic Solow growth model in continuous time with constant population growth and no technological change is globally asymptotically stable and, starting from any $k(0) > 0$, $k(t)$ monotonically converges to k^*.*

Proof. The proof of stability is now simpler and follows immediately from part 3 of Corollary 2.2 by noting that when $k < k^*$, we have $sf(k) - (n+\delta)k > 0$, and when $k > k^*$, we have $sf(k) - (n+\delta)k < 0$. ■

Figure 2.9 shows the analysis of stability diagrammatically. The figure plots the right-hand side of (2.33) and makes it clear that when $k < k^*$, $\dot{k} > 0$, and when $k > k^*$, $\dot{k} < 0$, so that the capital-labor ratio monotonically converges to the steady-state value k^*.

Example 2.2 (Dynamics with the Cobb-Douglas Production Function) *Let us return to the Cobb-Douglas production function introduced in Example 2.1:*

$$F(K, L, A) = AK^{\alpha}L^{1-\alpha}, \quad \text{with } 0 < \alpha < 1.$$

As noted above, the Cobb-Douglas production function is special, mainly because it has an elasticity of substitution between capital and labor equal to 1. For a homothetic *production function $F(K, L)$, the elasticity of substitution is defined by*

$$\sigma \equiv -\left[\frac{\partial \log(F_K/F_L)}{\partial \log(K/L)}\right]^{-1}, \tag{2.37}$$

where F_K and F_L denote the marginal products of capital and labor. (F is homothetic when F_K/F_L is only a function of K/L). For the Cobb-Douglas production function $F_K/F_L = \alpha L/((1-\alpha)K)$, and thus $\sigma = 1$. This feature implies that when the production function is Cobb-Douglas and factor markets are competitive, equilibrium factor shares will be constant regardless of the capital-labor ratio. In particular, the share of capital in national income is

$$\begin{aligned}\alpha_K(t) &= \frac{R(t)K(t)}{Y(t)} \\ &= \frac{F_K(K(t), L(t))K(t)}{Y(t)} \\ &= \frac{\alpha A K(t)^{\alpha-1}L(t)^{1-\alpha}K(t)}{AK(t)^{\alpha}L(t)^{1-\alpha}} \\ &= \alpha.\end{aligned}$$

Similarly, the share of labor is $\alpha_L(t) = 1-\alpha$. Thus with an elasticity of substitution equal to 1, as capital increases, its marginal product decreases proportionally, leaving the capital share (the amount of capital times its marginal product) constant.

Recall that with the Cobb-Douglas technology, the per capita production function takes the form $f(k) = Ak^{\alpha}$, so the steady state is again given by (2.34) (with population growth at the rate n) as

$$A(k^*)^{\alpha-1} = \frac{n+\delta}{s},$$

or as

$$k^* = \left(\frac{sA}{n+\delta}\right)^{\frac{1}{1-\alpha}},$$

which is a simple expression for the steady-state capital-labor ratio. It follows that k^* is increasing in s and A and decreasing in n and δ (these results are naturally consistent with those in Proposition 2.8). In addition, k^* is increasing in α, because a higher α implies less diminishing returns to capital, thus a higher capital-labor ratio is necessary to reduce the average return to capital to a level consistent with the steady state as given in (2.34).

Transitional dynamics are also straightforward in this case. In particular,

$$\dot{k}(t) = sAk(t)^{\alpha} - (n+\delta)k(t)$$

with initial condition $k(0) > 0$. To solve this equation, let $x(t) \equiv k(t)^{1-\alpha}$, so the equilibrium law of motion of the capital-labor ratio can be written in terms of $x(t)$ as

$$\dot{x}(t) = (1-\alpha)\, sA - (1-\alpha)(n+\delta)x(t),$$

which is a linear differential equation with a general solution

$$x(t) = \frac{sA}{n+\delta} + \left[x(0) - \frac{sA}{n+\delta}\right]\exp(-(1-\alpha)(n+\delta)t)$$

(see Appendix B). Expressing this solution in terms of the capital-labor ratio yields

$$k(t) = \left\{\frac{sA}{n+\delta} + \left[k(0)^{1-\alpha} - \frac{sA}{n+\delta}\right]\exp\left(-\,(1-\alpha)\,(n+\delta)\,t\right)\right\}^{\frac{1}{1-\alpha}}.$$

This solution illustrates that starting from any $k(0)$, the equilibrium $k(t) \to k^* = (sA/(n+\delta))^{1/(1-\alpha)}$, and in fact, the rate of adjustment is related to $(1-\alpha)(n+\delta)$. More specifically, the gap between $k(0)$ and the steady-state value k^* narrows at the exponential rate $(1-\alpha)(n+\delta)$. This result is intuitive: a higher α implies less diminishing returns to capital, which slows down the rate at which the marginal and average products of capital decline as capital accumulates, and this reduces the rate of adjustment to the steady state. Similarly, a smaller δ means less depreciation and a smaller n means slower population growth, both of which slow down the adjustment of capital per worker and thus the rate of transitional dynamics.

Example 2.3 (The Constant Elasticity of Substitution Production Function) *The Cobb-Douglas production function, which features an elasticity of substitution equal to 1, is a special case of the Constant Elasticity of Substitution (CES) production function, first introduced by Arrow et al. (1961). This production function imposes a constant elasticity, σ, not necessarily equal to 1. Consider a vector-valued index of technology $\mathbf{A}(t) = \left(A_H(t), A_K(t), A_L(t)\right)$. Then the CES production function can be written as*

$$\begin{aligned} Y(t) &= F(K(t), L(t), \mathbf{A}(t)) \\ &\equiv A_H(t)\left[\gamma\left(A_K(t)K(t)\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)(A_L(t)L(t))^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \end{aligned} \tag{2.38}$$

where $A_H(t) > 0$, $A_K(t) > 0$ and $A_L(t) > 0$ are three different types of technological change that are discussed further in Section 2.7; $\gamma \in (0, 1)$ is a distribution parameter, which determines how important labor and capital services are in determining the production of the final good; and $\sigma \in [0, \infty]$ is the elasticity of substitution. To verify that σ is the constant elasticity of substitution, let us use (2.37). The ratio of the marginal product of capital to the marginal productive labor, F_K/F_L, is then given by

$$\frac{F_K}{F_L} = \frac{\gamma A_K(t)^{\frac{\sigma-1}{\sigma}} K(t)^{-\frac{1}{\sigma}}}{(1-\gamma)\, A_L(t)^{\frac{\sigma-1}{\sigma}} L(t)^{-\frac{1}{\sigma}}},$$

so that the elasticity of substitution is indeed given by σ, that is,

$$\sigma = -\left[\frac{\partial \log(F_K/F_L)}{\partial \log(K/L)}\right]^{-1}.$$

The CES production function is particularly useful because it is more general and flexible than the Cobb-Douglas form while still being tractable. As $\sigma \to 1$, the CES production function (2.38) converges to the Cobb-Douglas function

$$Y(t) = A_H(t)(A_K(t))^{\gamma}(A_L(t))^{1-\gamma}(K(t))^{\gamma}(L(t))^{1-\gamma}.$$

As $\sigma \to \infty$, the CES production function becomes linear, that is,

$$Y(t) = \gamma A_H(t)A_K(t)K(t) + (1-\gamma)\, A_H(t)A_L(t)L(t).$$

Finally, as $\sigma \to 0$, the CES production function converges to the Leontief production function with no substitution between factors:

$$Y(t) = A_H(t) \min\left\{\gamma A_K(t)K(t); (1-\gamma)\, A_L(t)L(t)\right\}.$$

The special feature of the Leontief production function is that if $\gamma A_K(t)K(t) \neq (1-\gamma)\, A_L(t)L(t)$, either capital or labor will be partially idle in the sense that a small reduction in capital or labor will have no effect on output or factor prices. Exercise 2.23 illustrates a number of the properties of the CES production function, while Exercise 2.16 provides an alternative derivation of this production function along the lines of the original article by Arrow et al. (1961). Notice also that the CES production function with $\sigma > 1$ violates Assumption 1 (see Exercise 2.24), so in the context of aggregate production functions with capital and labor, we may take $\sigma \leq 1$ as the benchmark.

2.6 A First Look at Sustained Growth

Can the Solow model generate sustained growth without technological progress? The answer is yes, but only if some of the assumptions imposed so far are relaxed. The Cobb-Douglas example (Example 2.2) already showed that when α is close to 1, the adjustment of the capital-labor ratio back to its steady-state level can be very slow. A very slow adjustment toward a steady state has the flavor of sustained growth rather than the economy (quickly) settling down to a steady state. In fact, the simplest model of sustained growth essentially takes $\alpha = 1$ in terms of the Cobb-Douglas production function. To construct such a model, let us relax Assumptions 1 and 2 (which do not allow $\alpha = 1$), and consider the so-called AK model, where

$$F(K(t), L(t), A(t)) = AK(t) \tag{2.39}$$

and $A > 0$ is a constant. The results here apply with more general constant returns to scale production functions that relax Assumption 2, for example, with

$$F(K(t), L(t), A(t)) = AK(t) + BL(t). \tag{2.40}$$

Nevertheless, it is simpler to illustrate the main insights with (2.39), leaving the analysis of the case when the production function is given by (2.40) to Exercise 2.22.

Let us continue to assume that population grows at a constant rate n as before (see (2.32)). Then combining (2.32) with the production function (2.39), the fundamental law of motion of the capital stock becomes

$$\frac{\dot{k}(t)}{k(t)} = sA - \delta - n.$$

This equation shows that when the parameters satisfy the inequality $sA - \delta - n > 0$, there will be sustained growth in the capital-labor ratio and thus in output per capita. This result is summarized in the next proposition.

Proposition 2.10 *Consider the Solow growth model with the production function (2.39) and suppose that $sA - \delta - n > 0$. Then in equilibrium there is sustained growth of output per capita at the rate $sA - \delta - n$. In particular, starting with a capital-labor ratio $k(0) > 0$, the economy has*

$$k(t) = \exp\left((sA - \delta - n)\, t\right) k(0)$$

and

$$y(t) = \exp\left((sA - \delta - n)\, t\right) Ak(0).$$

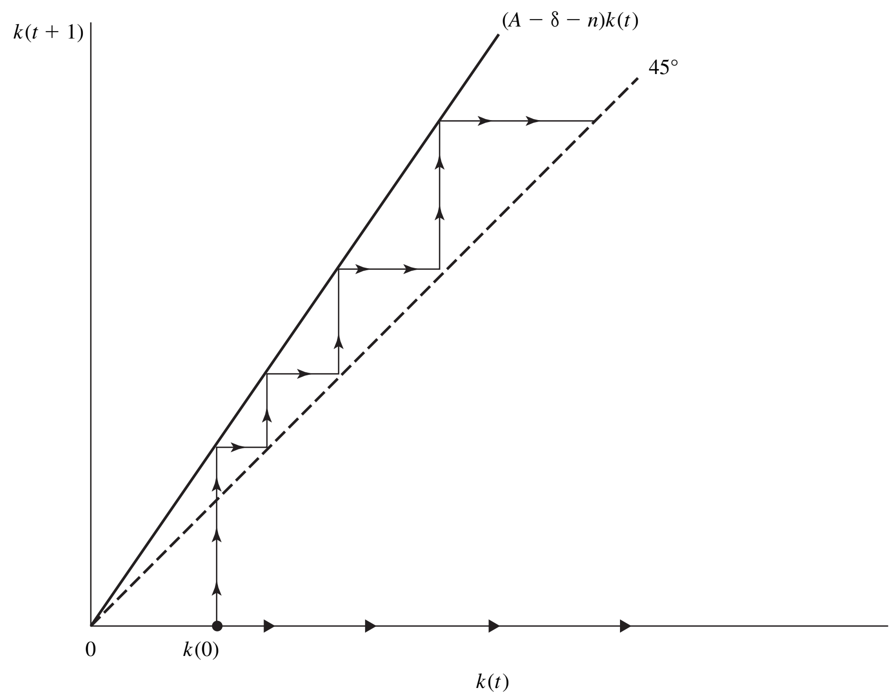

FIGURE 2.10 Sustained growth with the linear AK technology with $sA - \delta - n > 0$.

This proposition not only establishes the possibility of sustained growth but also shows that when the aggregate production function is given by (2.39), sustained growth is achieved without transitional dynamics. The economy always grows at a constant rate $sA - \delta - n$, regardless of the initial level of capital-labor ratio. Figure 2.10 shows this equilibrium diagrammatically.

Does the AK model provide an appealing approach to explaining sustained growth? While its simplicity is a plus, the model has a number of unattractive features. First, it is somewhat of a knife-edge case, which does not satisfy Assumptions 1 and 2; in particular, it requires the production function to be ultimately linear in the capital stock. Second and relatedly, this feature implies that as time goes by the share of national income accruing to capital will increase toward 1 (if it is not equal to 1 to start with). The next section shows that this tendency does not seem to be borne out by the data. Finally and most importantly, a variety of evidence suggests that technological progress is a major (perhaps the most significant) factor in understanding the process of economic growth. A model of sustained growth without technological progress fails to capture this essential aspect of economic growth. Motivated by these considerations, we next turn to the task of introducing technological progress into the baseline Solow growth model.

2.7 Solow Model with Technological Progress

2.7.1 Balanced Growth

The models analyzed so far did not feature technological progress. I now introduce changes in $A(t)$ to capture improvements in the technological knowhow of the economy. There is little doubt that today human societies know how to produce many more goods and can do so more efficiently than in the past. The productive knowledge of human society has

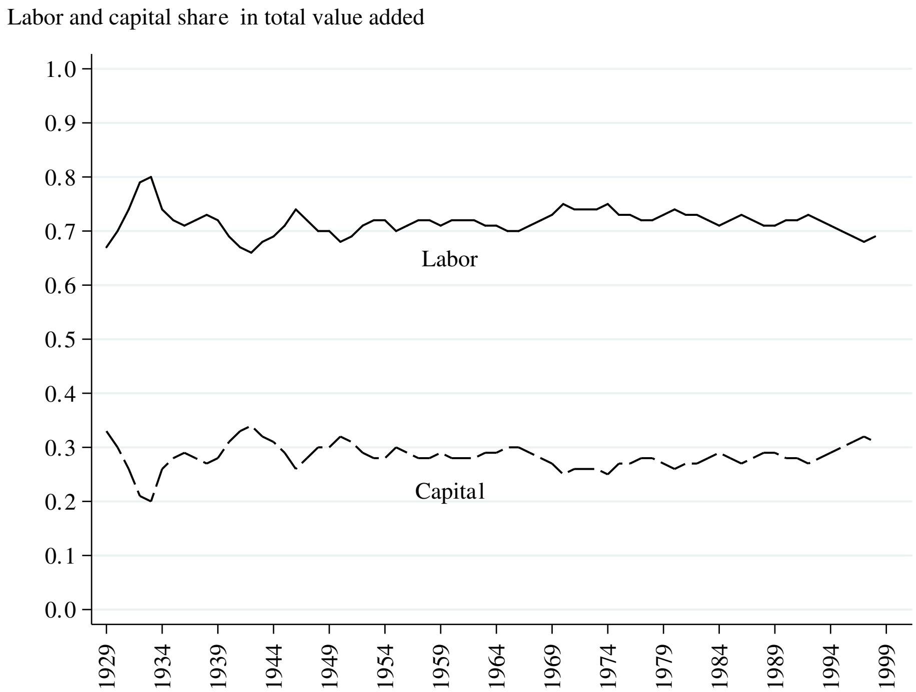

FIGURE 2.11 Capital and labor share in the U.S. GDP.

progressed tremendously over the past 200 years, and even more so over the past 1,000 or 10,000 years. This suggests that an attractive way of introducing economic growth in the framework developed so far is to allow technological progress in the form of changes in $A(t)$.

The key question is how to model the effects of changes in $A(t)$ on the aggregate production function. The standard approach is to impose discipline on the form of technological progress (and its impact on the aggregate production function) by requiring that the resulting allocations be consistent with balanced growth, as defined by the so-called Kaldor facts (Kaldor, 1963). Kaldor noted that while output per capita increases, the capital-output ratio, the interest rate, and the distribution of income between capital and labor remain roughly constant. Figure 2.11, for example, shows the evolution of the shares of capital and labor in the U.S. national income. Throughout the book, balanced growth refers to an allocation where output grows at a constant rate and capital-output ratio, the interest rate, and factor shares remain constant. (Clearly, three of these four features imply the fourth.)

Figure 2.11 shows that, despite fairly large fluctuations, there is no trend in factor shares. Moreover, a range of evidence suggests that there is no apparent trend in interest rates over long time horizons (see, e.g., Homer and Sylla, 1991). These facts and the relative constancy of capital-output ratios until the 1970s make many economists prefer models with balanced growth to those without. The share of capital in national income and the capital-output ratio are not exactly constant. For example, since the 1970s both the share of capital in national income and the capital-output ratio may have increased, depending on how one measures them. Nevertheless, constant factor shares and a constant capital-output ratio provide a good approximation to reality and a very useful starting point for our models.

Also for future reference, note that in Figure 2.11 the capital share in national income is about 1/3, while the labor share is about 2/3. This estimate ignores the share of land; land is not a major factor of production in modern economies (though this has not been true historically and is not true for the less-developed economies of today). Exercise 2.11 illustrates how incorporating land into this framework changes the analysis. Note also that this pattern of the factor distribution of income, combined with economists' desire to work with simple models, often makes them choose a Cobb-Douglas aggregate production function of the form $AK^{1/3}L^{2/3}$ as an approximation to reality (especially since it ensures that factor shares are constant by construction). However, Theorem 2.6 below shows that Cobb-Douglas technology is not necessary for balanced growth, and as noted in Example 2.2, the Cobb-Douglas form is both special and restrictive.

Another major advantage of models with balanced growth is that they are much easier to analyze than those with nonbalanced growth. Analysis is facilitated because with balanced growth, the equations describing the law of motion of the economy can be represented by difference or differential equations with well-defined steady states in transformed variables (thus, balanced growth will imply $\dot{k} = 0$, except that now the definition of k is different). This enables us to apply the tools used in the analysis of stationary models to study economies with sustained growth. It is nonetheless important to bear in mind that in reality, growth has many nonbalanced features. For example, the share of different sectors changes systematically over the growth process, with agriculture shrinking and manufacturing first increasing and then shrinking. Ultimately, we would like to develop models that combine certain balanced features with these types of structural transformations. I return to these issues in Part VII of the book.

2.7.2 Types of Neutral Technological Progress

What types of restrictions does balanced growth place on our models? It turns out that the answer to this question is "quite a few." The production function $F(K(t), L(t), A(t))$ is too general to achieve balanced growth, and only some very special types of production functions are consistent with balanced growth. To develop this point, consider an aggregate production function $\tilde{F}$ and let us define different types of neutral technological progress. A first possibility is

$$\tilde{F}(K(t), L(t), A(t)) = A(t)F(K(t), L(t))$$

for some constant returns to scale function F. This functional form implies that the technology term $A(t)$ is simply a multiplicative constant in front of another (quasi-) production function F. This type of technological progress is referred to as "Hicks-neutral" after the famous British economist John Hicks. Figure 2.12 illustrates this form of technological progress by plotting the isoquants of the function $\tilde{F}(K(t), L(t), A(t))$, which correspond to combinations of labor and capital for a given technology $A(t)$ such that the level of production is constant. Hicks-neutral technological progress, in the first panel, corresponds to a relabeling of the isoquants (without any change in their shape).

Another alternative is to have capital-augmenting or Solow-neutral technological progress in the form

$$\tilde{F}(K(t), L(t), A(t)) = F(A(t)K(t), L(t)),$$

which is also referred to as "capital-augmenting progress," because a higher $A(t)$ is equivalent to the economy having more capital. This type of technological progress corresponds to the isoquants shifting inward as if the capital axis were being shrunk (since a higher A now corresponds to a greater level of effective capital). This type of progress is shown in panel B of Figure 2.12 for a doubling of $A(t)$.

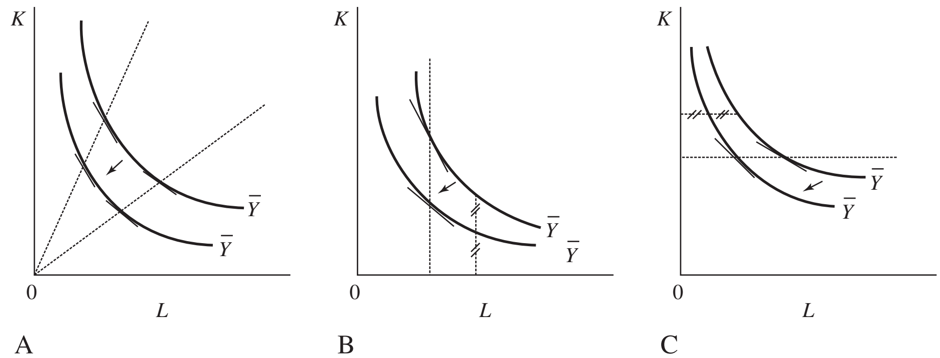

FIGURE 2.12 (A) Hicks-neutral, (B) Solow-neutral, and (C) Harrod-neutral shifts in isoquants.

Finally, we can have labor-augmenting or Harrod-neutral technological progress (panel C), named after Roy Harrod (whom we already encountered in the context of the Harrod-Domar model):

$$\tilde{F}(K(t), L(t), A(t)) = F\,[K(t), A(t)L(t)].$$

This functional form implies that an increase in technology $A(t)$ increases output as if the economy had more labor and thus corresponds to an inward shift of the isoquant as if the labor axis were being shrunk. The approximate form of the shifts in the isoquants are plotted in the third panel of Figure 2.12, again for a doubling of $A(t)$.

Of course in practice technological change can be a mixture of these, so we could have a vector-valued index of technology $\mathbf{A}(t) = \left(A_H(t), A_K(t), A_L(t)\right)$ and a production function that looks like

$$\tilde{F}(K(t), L(t), \mathbf{A}(t)) = A_H(t) F\left[A_K(t)K(t), A_L(t)L(t)\right], \tag{2.41}$$

which nests the CES production function introduced in Example 2.3. Nevertheless, even (2.41) is a restriction on the form of technological progress, since in general changes in technology, $A(t)$, could modify the entire production function.

Although all of these forms of technological progress look equally plausible ex ante, we will next see that balanced growth in the long run is only possible if all technological progress is labor-augmenting or Harrod-neutral. This result is very surprising and troubling, since there are no compelling reasons for why technological progress should take this form. I return to a discussion of why long-run technological change might be Harrod-neutral in Chapter 15.

2.7.3 Uzawa's Theorem

The discussion above suggests that the key elements of balanced growth are the constancy of factor shares and the constancy of the capital-output ratio, $K(t)/Y(t)$. The shares of capital and labor in national income are

$$\alpha_K(t) \equiv \frac{R(t)K(t)}{Y(t)} \quad \text{and} \quad \alpha_L(t) \equiv \frac{w(t)L(t)}{Y(t)}.$$

By Assumption 1 and Theorem 2.1, $\alpha_K(t) + \alpha_L(t) = 1$.

A version of the following theorem was first proved by the leading growth theorist, Hirofumi Uzawa (1961). The statement and the proof here build on the argument developed in the recent paper by Schlicht (2006). The theorem shows that constant growth of output, capital, and consumption combined with constant returns to scale implies that the aggregate production function must have a representation with Harrod-neutral (purely labor-augmenting) technological progress. For simplicity and without loss of any generality, I focus on continuous-time models.

Theorem 2.6 (Uzawa's Theorem I) *Consider a growth model with aggregate production function*

$$Y(t) = \tilde{F}(K(t), L(t), \tilde{A}(t)),$$

where $\tilde{F} : \mathbb{R}^2_+ \times \mathcal{A} \to \mathbb{R}_+$ *and* $\tilde{A}(t) \in \mathcal{A}$ *represents technology at time* t *(where* $\mathcal{A}$ *is an arbitrary set, e.g., a subset of* $\mathbb{R}^N$ *for some natural number* N*). Suppose that* $\tilde{F}$ *exhibits constant returns to scale in* K *and* L*. The aggregate resource constraint is*

$$\dot{K}(t) = Y(t) - C(t) - \delta K(t).$$

Suppose that there is a constant growth rate of population, $L(t) = \exp(nt)L(0)$*, and that there exists* $T < \infty$ *such that for all* $t \geq T$*,* $\dot{Y}(t)/Y(t) = g_Y > 0$*,* $\dot{K}(t)/K(t) = g_K > 0$*, and* $\dot{C}(t)/C(t) = g_C > 0$*. Then*

1. $g_Y = g_K = g_C$*; and*
2. *for any* $t \geq T$*, there exists a function* $F : \mathbb{R}^2_+ \to \mathbb{R}_+$ *homogeneous of degree 1 in its two arguments, such that the aggregate production function can be represented as*

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

where $A(t) \in \mathbb{R}_+$ *and*

$$\frac{\dot{A}(t)}{A(t)} = g = g_Y - n.$$

Proof. (Part 1) By hypothesis for $t \geq T$, we have $Y(t) = \exp(g_Y(t-T))Y(T)$, $K(t) = \exp(g_K(t-T))K(T)$, and $L(t) = \exp(n(t-T))\,L(T)$. Since $\dot{K}(t) = g_K K(t)$, the aggregate resource constraint at time t implies

$$(g_K + \delta)K(t) = Y(t) - C(t).$$

Dividing both sides of this equation by $\exp\left(g_K(t-T)\right)$, we obtain

$$(g_K + \delta)K(T) = \exp((g_Y - g_K)(t-T))Y(T) - \exp((g_C - g_K)(t-T))C(T)$$

for all $t \geq T$. Differentiating the previous equation with respect to time yields

$$(g_Y - g_K)\exp((g_Y - g_K)(t-T))Y(T) - (g_C - g_K)\exp((g_C - g_K)(t-T))C(T) = 0$$

for all $t \geq T$. This equation can hold for all t if any of the following four conditions is true: (i) $g_Y = g_K = g_C$, (ii) $g_Y = g_C$ and $Y(T) = C(T)$, (iii) if $g_Y = g_K$ and $C(T) = 0$, or (iv) $g_C = g_K$ and $Y(T) = 0$. The latter three possibilities contradict, respectively, that $g_K > 0$, that $g_C > 0$ (which implies $C(T) > 0$ and $K(T) > 0$, and hence $Y(T) > C(T)$), and that

$Y(T) > 0$. Therefore (i) must apply and $g_Y = g_K = g_C$, as claimed in the first part of the theorem.

(Part 2) For any $t \geq T$, the aggregate production function for time T can be written as

$$\exp\left(-g_Y\,(t-T)\right) Y(t) = \tilde{F}\left[\exp\left(-g_K\,(t-T)\right) K(t),\, \exp\left(-n(t-T)\right) L(t),\, \tilde{A}(T)\right].$$

Multiplying both sides by $\exp\left(g_Y(t-T)\right)$ and using the constant returns to scale property of $\tilde{F}$ yields

$$Y(t) = \tilde{F}\left[\exp\left((t-T)\left(g_Y - g_K\right)\right) K(t),\, \exp\left((t-T)(g_Y - n)\right) L(t),\, \tilde{A}(T)\right].$$

From part 1, $g_Y = g_K$ and thus for any $t \geq T$,

$$Y(t) = \tilde{F}\left[K(t),\, \exp\left((t-T)\left(g_Y - n\right)\right) L(t),\, \tilde{A}(T)\right]. \tag{2.42}$$

Since (2.42) is true for all $t \geq T$ and $\tilde{F}$ is homogeneous of degree 1 in K and L, there exists a function $F : \mathbb{R}^2_+ \to \mathbb{R}_+$ that is homogeneous of degree 1 such that

$$Y(t) = F(K(t), \exp((g_Y - n)t)L(t)).$$

Rewriting this as

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

with

$$\frac{\dot{A}(t)}{A(t)} = g_Y - n$$

establishes the second part of the theorem. ■

A remarkable feature of this theorem is that it was stated and proved without any reference to equilibrium behavior. The theorem only exploits the fact that the production function exhibits constant returns to scale in K and L and that the allocation $[Y(t), K(t), C(t)]_{t=0}^{\infty}$ features output, capital, and consumption all growing at the same constant rate after time T. Notice, however, that the theorem holds under the hypothesis that $Y(t)$, $K(t)$, and $C(t)$ exhibit constant growth rates after some (finite) time T. A stronger result would require the same conclusions to hold as $t \to \infty$. Exercise 2.14 contains a generalization of Theorem 2.6 in this direction, but also shows why some additional conditions need to be imposed in this case.

Before providing an economic intuition for Theorem 2.6, let us state a simple corollary of this theorem, which will be useful both in the discussions below and for the intuition.

Corollary 2.3 *Under the assumptions of Theorem 2.6, for all $t \geq T$ technological progress can be represented as Harrod neutral (purely labor-augmenting).*

In light of Theorem 2.6 and this corollary, with a slight abuse of terminology, we can say that "technological change must be asymptotically Harrod neutral."

Let us now return to the intuition for Uzawa's Theorem. This theorem supposes that there is capital accumulation, that is, $g_K > 0$. Part 1 implies that this is possible only if output and capital grow at the same rate. Either this growth rate is equal to the rate of population growth, n, in which case there is no technological change (the proposition applies with $g_Y = 0$) or the economy exhibits growth of per capita income and the capital-labor ratio ($g_K = g_Y > 0$).

The latter case creates an asymmetry between capital and labor, in the sense that capital is accumulating faster than labor. Constancy of growth then requires technological change to make up for this asymmetry—that is, technology should take a labor-augmenting form.

This intuition does not provide a reason for why technology should take this labor-augmenting (Harrod-neutral) form, however. The theorem and its corollary simply state that if technology did not take this form, an (asymptotic) allocation with constant growth rates of output, capital and consumption (and thus balanced growth) would not be possible. At some level, this result is distressing, since it implies that balanced growth (in fact something weaker than balanced growth) is only possible under a very stringent assumption. Chapter 15 shows that when technology is endogenous the same intuition also implies that technology should be endogenously more labor-augmenting than capital-augmenting.

Notice also that Theorem 2.6 and its corollary do *not* state that technological change has to be labor-augmenting all the time. Instead, technological change ought to be labor-augmenting after time T (along the balanced growth path). This is the pattern that certain classes of endogenous technology models will generate (see again Chapter 15 for a discussion). More importantly, contrary to common claims in textbooks and the literature, Theorem 2.6 does *not* even state that capital-augmenting (Solow-neutral) technological change is impossible as $t \to \infty$. It states that such technological progress is not possible if there is balanced growth after some date T. Exercise 2.17 provides a simple example where asymptotic balanced growth (with the conditions in Theorem 2.6 being satisfied as $t \to \infty$) is possible in the presence of asymptotic capital-augmenting technological progress.

It should also be emphasized that Theorem 2.6 does not require that $Y(t) = F(K(t), A(t)L(t))$, but only that it has a representation of the form $Y(t) = F(K(t), A(t)L(t))$. For example, if the aggregate production function is Cobb-Douglas, that is,

$$Y(t) = (A_K(t)K(t))^{\alpha}(A_L(t)L(t))^{1-\alpha},$$

then both $A_K(t)$ and $A_L(t)$ could grow at constant rates while maintaining balanced growth. However, in this Cobb-Douglas example we can define $A(t) = A_K(t)^{\alpha/(1-\alpha)}A_L(t)$ and the production function can be represented as

$$Y(t) = K(t)^{\alpha}\ (A(t)L(t))^{1-\alpha},$$

so that technological change is represented as purely labor-augmenting, which is what Theorem 2.6 requires. Intuitively, the differences between labor-augmenting and capital-augmenting (and Hicks-neutral) forms of technological progress matter when the elasticity of substitution between capital and labor is not equal to 1. In the Cobb-Douglas case, as we have seen above, this elasticity of substitution is equal to 1; thus Harrod-neutral, Solow-neutral, and Hicks-neutral forms of technological progress are simple transforms of one another.

Theorem 2.6 does not specify how factor prices behave. As noted at the beginning of this section, the Kaldor facts also require constant factor shares. Since capital and output are growing at the same rate, the rental rate of capital must be constant. Does Theorem 2.6 (combined with competitive factor markets) imply constant factor shares? Unfortunately, the answer is not necessarily. This is related to an implicit limitation in Theorem 2.6. The theorem states that the original production function $\tilde{F}(K(t), L(t), \tilde{A}(t))$ has a representation of the form $F(K(t), A(t)L(t))$ along an asymptotic path with constant growth rates. This does not guarantee that the derivatives of $\tilde{F}$ and F with respect to K and L agree. Exercise 2.19 provides an example of production function $\tilde{F}$ that satisfies all of the conditions of Theorem 2.6 (and thus admits a representation of the form $F(K(t), A(t)L(t))$ as $t \to \infty$), but has derivatives that do not agree with those of F. In fact, the exercise shows that, with competitive markets, this $\tilde{F}$

leads to arbitrary behavior of factor prices as $t \to \infty$. The next theorem, however, shows that along a balanced growth path, where factor shares are also constant over time, the derivatives of $\tilde{F}$ and F must agree, and vice versa.

Theorem 2.7 (Uzawa's Theorem II) *Suppose that all hypotheses in Theorem 2.6 are satisfied, so that $\tilde{F} : \mathbb{R}^2_+ \times \mathcal{A} \to \mathbb{R}_+$ has a representation of the form $F(K(t), A(t)L(t))$ with $A(t) \in \mathbb{R}_+$ and $\dot{A}(t)/A(t) = g = g_Y - n$ (for $t \geq T$). In addition, suppose that factor markets are competitive and that for all $t \geq T$, the rental rate satisfies $R(t) = R^*$ (or equivalently, $\alpha_K(t) = \alpha^*_K$). Then, denoting the partial derivatives of $\tilde{F}$ and F with respect to their first two arguments by $\tilde{F}_K$, $\tilde{F}_L$, F_K, and F_L, we have*

$$\begin{aligned} \tilde{F}_K(K(t), L(t), \tilde{A}(t)) &= F_K(K(t), A(t)L(t)) \text{ and} \\ \tilde{F}_L(K(t), L(t), \tilde{A}(t)) &= A(t)F_L(K(t), A(t)L(t)). \end{aligned} \tag{2.43}$$

Moreover, if (2.43) holds and factor markets are competitive, then $R(t) = R^*$ (and $\alpha_K(t) = \alpha^*_K$) for all $t \geq T$.

Proof. From Theorem 2.6, $g_Y = g_K = g_C = g + n$. Since $R(t) = R^*$ for all $t \geq T$, this also implies that the wage rate satisfies $w(t) = \left(Y(t) - R^*K(t)\right)/L(t) = \exp\left(g(t-T)\right) w^*$ (where $w^* = w(T)$). Therefore we have that for all $t \geq T$,

$$\begin{aligned} R^* &= \tilde{F}_K(K(t), L(t), \tilde{A}(t)) \\ \exp\left(g(t-T)\right) w^* &= \tilde{F}_L(K(t), L(t), \tilde{A}(t)). \end{aligned} \tag{2.44}$$

With the same argument as in the proof of Theorem 2.6, we can also write

$$\begin{aligned} R^* &= \tilde{F}_K(\exp\left(-(g+n)\,(t-T)\right) K(t), \exp\left(-n(t-T)\right) L(t), \tilde{A}(T)), \\ w^* &= \tilde{F}_L(\exp\left(-(g+n)(t-T)\right) K(t), \exp\left(-n(t-T)\right) L(t), \tilde{A}(T)). \end{aligned}$$

Using the fact that $\tilde{F}_K$ and $\tilde{F}_L$ are homogeneous of degree 0 in K and L (see Theorem 2.1), the previous two equations can be rewritten as

$$\begin{aligned} R^* &= \tilde{F}_K(K(t), \exp\left(g(t-T)\right) L(t), \tilde{A}(T)), \\ w^* &= \tilde{F}_L(K(t), \exp\left(g(t-T)\right) L(t), \tilde{A}(T)). \end{aligned}$$

Comparing these to (2.44), we can conclude that for all $t \geq T$,

$$\begin{aligned} \tilde{F}_K(K(t), \exp\left(g(t-T)\right) L(t), \tilde{A}(T)) &= \tilde{F}_K(K(t), L(t), \tilde{A}(t)), \\ \exp\left(g(t-T)\right) \tilde{F}_L(K(t), \exp\left(g(t-T)\right) L(t), \tilde{A}(T)) &= \tilde{F}_L(K(t), L(t), \tilde{A}(t)). \end{aligned}$$

This implies that there exist functions homogeneous of degree 0, $\hat{F}_1, \hat{F}_2 : \mathbb{R}^2_+ \to \mathbb{R}_+$, such that

$$\begin{aligned} \hat{F}_1(K(t), A(t)L(t)) &= \tilde{F}_K(K(t), L(t), \tilde{A}(t)), \\ A(t)\hat{F}_2(K(t), A(t)L(t)) &= \tilde{F}_L(K(t), L(t), \tilde{A}(t)), \end{aligned}$$

with $\dot{A}(t)/A(t) = g$. Define $\hat{F} : \mathbb{R}_+^2 \to \mathbb{R}_+$ as

$$\hat{F}(K, AL) \equiv \hat{F}_1(K, AL)K + \hat{F}_2(K, AL)AL. \tag{2.45}$$

From Theorem 2.1, $\hat{F}(K(t), A(t)L(t)) = \tilde{F}(K(t), L(t), \tilde{A}(t))$, and thus $\hat{F}$ is homogeneous of degree 1 in its two arguments and provides a representation of $\tilde{F}$ along the path $[K(t), L(t)]_{t=0}^{\infty}$. Since $\hat{F}$ is homogeneous of degree 1, (2.45) implies that its partial derivatives are given by $\hat{F}_1$ and $\hat{F}_2$ and thus agree with those of $\tilde{F}$, establishing (2.43).

To prove the second part of the theorem, simply note that with competitive factor markets, we have that for $t \geq T$,

$$\begin{aligned}\alpha_K(t) &\equiv \frac{R(t)K(t)}{Y(t)} \\ &= \frac{K(t)}{Y(t)} \frac{\partial \tilde{F}(K(t), L(t), \tilde{A}(t))}{\partial K(t)} \\ &= \alpha_K^*,\end{aligned}$$

where the second line uses the definition of the rental rate of capital in a competitive market, and the third line uses (2.43) together with the fact that F is homogeneous of degree 1. ■

Theorem 2.7 implies that any allocation with constant growth rates for output, capital, and consumption must be a balanced growth path (where factor shares in national income are also constant). It also implies that balanced growth can only be generated by an aggregate production function that features Harrod-neutral technological change.

A further intuition for Theorem 2.6 comes from Theorem 2.7. Suppose the production function takes the special form $F(A_K(t)K(t), A_L(t)L(t))$. Theorem 2.7 implies that factor shares must be constant as $t \to \infty$. Thus, given constant returns to scale, balanced growth after some time T is possible only when total capital inputs, $A_K(t)K(t)$, and total labor inputs, $A_L(t)L(t)$, grow at the same rate; otherwise, the share of either capital or labor will not be constant. But if total labor and capital inputs grow at the same rate, then output $Y(t)$ must also grow at this rate (again because of constant returns to scale). The fact that the capital-output ratio is constant in steady state then implies that $K(t)$ must grow at the same rate as output and thus at the same rate as $A_L(t)L(t)$. Therefore, balanced growth is only possible if $A_K(t)$ is constant after date T.

2.7.4 The Solow Growth Model with Technological Progress: Continuous Time

I now present an analysis of the Solow growth model with technological progress in continuous time. The discrete-time case can be analyzed analogously, and I omit the details to avoid repetition. Theorem 2.6 implies that when the economy is experiencing balanced growth, the production function must have a representation of the form

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

with purely labor-augmenting technological progress. Most macroeconomic and growth analyses then assume that it takes this form throughout (for all t) and that there is technological progress at the rate $g > 0$, that is,

$$\frac{\dot{A}(t)}{A(t)} = g > 0. \tag{2.46}$$

Let us also start with this assumption. Suppose also that population grows at the rate n as in (2.32). Again using the constant saving rate, capital accumulates according to the differential equation

$$\dot{K}(t) = sF(K(t), A(t)L(t)) - \delta K(t). \tag{2.47}$$

The simplest way of analyzing this economy is to express everything in terms of a normalized variable. Since "effective" or efficiency units of labor are given by $A(t)L(t)$, and F exhibits constant returns to scale in its two arguments, I now define $k(t)$ as the *effective capital-labor* ratio (capital divided by efficiency units of labor) so that

$$k(t) \equiv \frac{K(t)}{A(t)L(t)}. \tag{2.48}$$

Although there is a slight danger that the use of the same symbol, $k(t)$, for capital-labor ratio earlier and effective capital-labor ratio now might cause some confusion, the important parallel between the roles of capital-labor ratio in the Solow model without technological progress and of the effective capital-labor ratio with labor-augmenting technological progress justifies this notation.

Differentiating this expression with respect to time, we obtain

$$\frac{\dot{k}(t)}{k(t)} = \frac{\dot{K}(t)}{K(t)} - g - n. \tag{2.49}$$

The quantity of output per unit of effective labor can be written as

$$\begin{aligned} \hat{y}(t) &\equiv \frac{Y(t)}{A(t)L(t)} \\ &= F\left(\frac{K(t)}{A(t)L(t)}, 1\right) \\ &\equiv f(k(t)). \end{aligned}$$

Income per capita is $y(t) \equiv Y(t)/L(t)$, so that

$$\begin{aligned} y(t) &= A(t)\hat{y}(t) \\ &= A(t)f(k(t)). \end{aligned} \tag{2.50}$$

It should be clear that if $\hat{y}(t)$ is constant, income per capita, $y(t)$, will grow over time, since $A(t)$ is growing. This highlights that in this model, and more generally in models with technological progress, we should not look for a steady state where income per capita is constant, but for a *balanced growth path* (BGP), where income per capita grows at a constant rate, while transformed variables, such as $\hat{y}(t)$ or $k(t)$ in (2.49), remain constant. Since these transformed variables remain constant, BGPs can be thought of as steady states of a transformed model. Motivated by this observation, in models with technological change throughout I use the terms "steady state" and "BGP" interchangeably. We will see that consistent with the definition in Section 2.7.1, this BGP allocation will also feature constant capital-output ratio, interest rate, and factor shares in national income.

Next, substituting for $\dot{K}(t)$ from (2.47) into (2.49),

$$\frac{\dot{k}(t)}{k(t)} = \frac{sF(K(t), A(t)L(t))}{K(t)} - (\delta + g + n).$$

Using (2.48),

$$\frac{\dot{k}(t)}{k(t)} = \frac{sf(k(t))}{k(t)} - (\delta + g + n), \tag{2.51}$$

which is very similar to the law of motion of the capital-labor ratio in the model without technological progress (2.33). The only difference is the presence of g, which reflects the fact that now k is no longer the capital-labor ratio, but the *effective* capital-labor ratio. Thus for k to remain constant in the BGP, the capital-labor ratio needs to increase at the rate g.

An equilibrium in this model is defined similarly to before. A steady state or a BGP is, in turn, defined as an equilibrium in which the effective capital-labor ratio $k(t)$ is constant. Consequently the following proposition holds (proof omitted).

Proposition 2.11 *Consider the basic Solow growth model in continuous time with Harrod-neutral technological progress at the rate g and population growth at the rate n. Suppose that Assumptions 1 and 2 hold, and define the effective capital-labor ratio as in (2.48). Then there exists a unique BGP where the effective capital-labor ratio is equal to $k^* \in (0, \infty)$ given by*

$$\frac{f(k^*)}{k^*} = \frac{\delta + g + n}{s}. \tag{2.52}$$

Per capita output and consumption grow at the rate g.

Equation (2.52), which determines the BGP (steady-state) effective capital-labor ratio, emphasizes that now total savings, $sf(k)$, are used for replenishing the capital stock for three distinct reasons. The first is again depreciation at the rate δ. The second is population growth at the rate n (which reduces capital per worker). The third is Harrod-neutral technological progress, which reduces effective capital-labor ratio at the rate g when the capital-labor ratio is constant. Thus the replenishment of the effective capital-labor ratio requires total investment to be equal to $(\delta + g + n)\,k$, which is the intuitive explanation for (2.52).

The comparative static results are also similar to before, with the additional comparative static with respect to the initial level of the labor-augmenting technology, $A(0)$ (the level of technology at all points in time, $A(t)$, is completely determined by $A(0)$ given the assumption in (2.46)).

Proposition 2.12 *Suppose Assumptions 1 and 2 hold and let $A(0)$ be the initial level of technology. Denote the BGP level of effective capital-labor ratio by $k^*(A(0), s, \delta, n, g)$ and the level of output per capita by $y^*(A(0), s, \delta, n, g, t)$ (the latter is a function of time, since it is growing over time). Then*

$$\frac{\partial k^*(A(0), s, \delta, n, g)}{\partial A(0)} = 0, \quad \frac{\partial k^*(A(0), s, \delta, n, g)}{\partial s} > 0,$$

$$\frac{\partial k^*(A(0), s, \delta, n, g)}{\partial n} < 0, \quad \text{and} \quad \frac{\partial k^*(A(0), s, \delta, n, g)}{\partial \delta} < 0,$$

and also

$$\frac{\partial y^*(A(0), s, \delta, n, g, t)}{\partial A(0)} > 0, \quad \frac{\partial y^*(A(0), s, \delta, n, g, t)}{\partial s} > 0,$$

$$\frac{\partial y^*(A(0), s, \delta, n, g, t)}{\partial n} < 0, \quad \text{and} \quad \frac{\partial y^*(A(0), s, \delta, n, g, t)}{\partial \delta} < 0,$$

for each t.

Proof. See Exercise 2.25. ■

Finally, the transitional dynamics of the economy with technological progress are similar to the dynamics without technological change.

Proposition 2.13 *Suppose that Assumptions 1 and 2 hold. Then the BGP of the Solow growth model with Harrod-neutral technological progress and population growth in continuous time is asymptotically stable; that is, starting from any $k(0) > 0$, the effective capital-labor ratio converges to the BGP value k^* ($k(t) \to k^*$).*

Proof. See Exercise 2.26. ■

Therefore, with Harrod-neutral technological change, the dynamics of the equilibrium path and the comparative statics are very similar to those in the model without technological progress. The major difference is that now the model generates growth in output per capita, so it can be mapped to the data more successfully. However, the disadvantage is that growth is driven entirely exogenously. The growth rate of the economy is exactly the same as the exogenous growth rate of the technology stock. The model specifies neither where this technology stock comes from nor how fast it grows.

2.8 Comparative Dynamics

This section briefly undertakes some simple comparative dynamics exercises. Comparative dynamics are different from the comparative static results in Propositions 2.3, 2.8, or 2.12 in that the focus is now on the entire path of adjustment of the economy following a shock or a change in parameters. The basic Solow model is particularly well suited to such an analysis because of its simplicity. These exercises are also useful because the basic Solow model and its neoclassical cousin are often used for analysis of policy changes, medium-run shocks, and business cycle dynamics, so an understanding of how the basic model responds to various shocks is useful in a range of applications.

Recall that the law of motion of the effective capital-labor ratio in the continuous-time Solow model is given by (2.51), that is, $\dot{k}(t)/k(t) = sf(k(t))/k(t) - (\delta + g + n)$. The right-hand side of this equation is plotted in Figure 2.13. The intersection with the horizontal axis gives the unique BGP, with effective capital-labor ratio k^*. This figure is sufficient for the analysis of comparative dynamics. Consider, for example, a one-time, unanticipated, permanent increase in the saving rate from s to s'. This shifts the curve to the right as shown by the dashed line, with a new intersection with the horizontal axis at k^{**}. The dashed arrows under the horizontal axis show how the effective capital-labor ratio adjusts gradually to the new BGP effective capital-labor ratio, k^{**}. Immediately after the increase in the saving rate is realized, the capital stock and the effective capital-labor ratio remain unchanged (since they are state variables). After this point, k follows the dashed arrows and converges monotonically to k^{**}. The

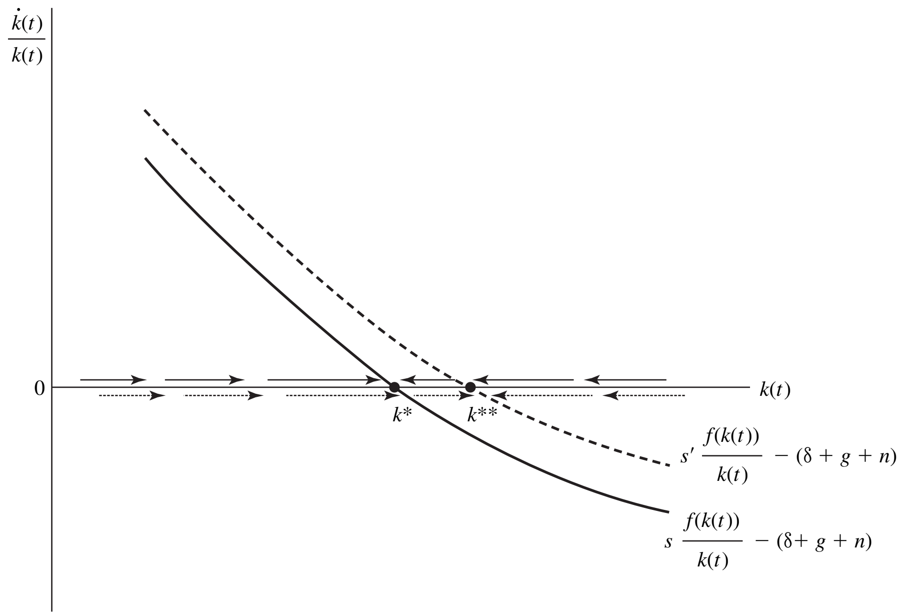

FIGURE 2.13 Dynamics following an increase in the saving rate from s to s'. The solid arrows show the dynamics for the initial steady state, while the dashed arrows show the dynamics for the new steady state.

comparative dynamics following a one-time, unanticipated, permanent decrease in δ or n are identical.

The same diagrammatic analysis can be used for studying the effect of an unanticipated but transitory change in parameters. For example, imagine that s changes in an unanticipated manner at $t = t'$, but this change will be reversed and the saving rate will return back to its original value at some known future date $t'' > t'$. In this case, starting at t', the economy follows the dashed arrows until t'. After t'', the original steady state of the differential equation applies and together with this, the solid arrows above the horizontal axis become effective. Thus from t'' onward, the economy gradually returns back to its original balanced growth equilibrium, k^*. We will see that similar comparative dynamics can be carried out in the neoclassical growth model as well, but the response of the economy to some of these changes will be more complex.

2.9 Taking Stock

What have we learned from the Solow model? At some level, a lot. We now have a simple and tractable framework that allows us to study capital accumulation and the implications of technological progress. As we will see in the next chapter, this framework is already quite useful in helping us think about the data.

However, at another level, we have learned relatively little. The questions that Chapter 1 posed are related to why some countries are rich while others are poor, why some countries grow while others stagnate, and why the world economy embarked upon the process of steady growth over the past few centuries. The Solow model shows us that if there is no technological progress, and as long as we are not in the AK world ruled out by Assumption 2, there will be no sustained growth. In this case, we can talk about cross-country output differences but not about growth of countries or of the world economy.

The Solow model generates per capita output growth only by introducing exogenous technological progress. But in this case, everything is driven by technological progress, and technological progress itself is exogenous, just a black box, outside the model and outside the influence of economic incentives. If technological progress is "where it's at," then we have to study and understand which factors generate technological progress, what makes some firms and societies invent better technologies, and what induces firms and societies to adopt and use these superior technologies.

Even on the question of capital accumulation, the Solow model is not entirely satisfactory. The rate of capital accumulation is determined by the saving rate, the depreciation rate, and the rate of population growth. All these rates are taken as exogenous.

In this light, the Solow growth model is most useful as a framework for laying out the general issues and questions. It emphasizes that to understand growth, we have to understand physical capital accumulation (and human capital accumulation, which is discussed in the next chapter) and perhaps most importantly, technological progress. All of these are black boxes in the Solow growth model. Therefore, much of the rest of the book will be about digging deeper, trying to uncover what lies in these black boxes. I start by introducing consumer optimization in Chapter 8, which enables a more systematic study of capital accumulation. Then I turn to models in which human capital accumulation and technological progress are endogenous. A model in which the rate of accumulation of factors of production and technology are endogenous gives us a framework for posing and answering questions related to the fundamental causes of economic growth.

Nevertheless, even in its bare-bones form, the Solow model is useful in helping us think about the world and bringing useful perspectives, especially related to the proximate causes of economic growth. This is the topic of the next chapter.

2.10 References and Literature

The model analyzed in this chapter was first developed in Solow (1956) and Swan (1956). Solow (1970) gives a nice and accessible treatment, with historical references. Barro and Sala-i-Martin's (2004, Chapter 1) textbook presents a more up-to-date treatment of the basic Solow model at the graduate level, while Jones (1998, Chapter 2) presents an excellent undergraduate treatment.

The treatment in the chapter made frequent references to basic consumer and general equilibrium theory. These are prerequisites for an adequate understanding of the theory of economic growth. Some of the important results from dynamic general equilibrium theory are discussed in Chapter 5. Mas-Colell, Whinston, and Green's (1995) graduate microeconomics textbook contains an excellent treatment of most of the necessary material, including producer theory and an accessible presentation of the basic notions of general equilibrium theory, including a discussion of Arrow securities and the definition of Arrow-Debreu commodities.

Properties of homogeneous functions and Euler's Theorem can be found, for example, in Simon and Blume (1994, Chapter 20). The reader should be familiar with the Implicit Function Theorem and properties of concave and convex functions, which are used throughout the book. A review is given in Appendix A.

Appendix B provides an overview of solutions to differential and difference equations and a discussion of stability. Theorems 2.2, 2.3, 2.4, and 2.5 follow from the results presented there. In addition, the reader may want to consult Boyce and DiPrima (1977), Luenberger (1979), or Simon and Blume (1994) for various results on difference and differential equations. Knowledge of solutions to simple differential equations and stability properties of difference and differential equations at the level of Appendix B is assumed in the text. In addition, the material in Luenberger (1979) is particularly useful, since it contains a unified treatment of difference and differential equations. Galor (2005) gives an introduction to difference equations and discrete-time dynamical systems for economists.

The golden rule saving rate was introduced by Edmund Phelps (1966). It is called the "golden rule" rate with reference to the biblical golden rule "do unto others as you would have them do unto you" applied in an intergenerational setting—that is, presuming that those living and consuming at each different date form a different generation. While the golden rule saving rate is of historical interest and is useful for discussions of dynamic efficiency, it has no intrinsic optimality property, since it is not derived from well-defined preferences. Optimal savings are discussed in greater detail in Chapter 8.

The balanced growth facts were first noted by Kaldor (1963). Figure 2.11 uses data from Piketty and Saez (2003). Homer and Sylla (1991) discuss the history of interest rates over many centuries and across different societies; they show that there is no notable upward or downward trend in interest rate. Nevertheless, not all aspects of the economic growth process are balanced, and the nonbalanced nature of growth is discussed in detail in Part VII of the book, which also contains references to changes in the sectoral composition of output in the course of the growth process.

A simpler version of Theorem 2.6 was first proved by Uzawa (1961). There are various different proofs in the literature, though many are not fully rigorous. The proof given here is adapted from Schlicht (2006), which is also discussed in Jones and Scrimgeour (2006). A similar proof also appears in Wan (1971). Barro and Sala-i-Martin (2004, Chapter 1) also suggest a proof. Nevertheless, their argument is incomplete, since it assumes that technological change must be a combination of Harrod- and Solow-neutral technological change, which is rather restrictive and is not necessary for the proof. The theorem and the proof provided here are therefore more general and complete. There are also a variety of misconceptions about the implications of Theorem 2.6. Many textbooks claim that this theorem rules out (asymptotic) capital-augmenting technological progress (unless the production function is Cobb-Douglas). Exercise 2.17 shows that this claim is not true and balanced growth is possible even with asymptotic capital-augmenting technological progress (with non–Cobb-Douglas production functions). Theorem 2.6 holds when balanced growth applies after some finite time T or under additional conditions, as discussed in Exercise 2.14. Moreover, it is also important to emphasize, as I did in the text, that Theorem 2.6 only provides a representation for a particular path of capital and labor. Consequently, this representation cannot always be used for equilibrium analysis (or for pricing capital and labor), as shown by Exercise 2.19. Theorem 2.7 was provided as a way of overcoming this difficulty. I am not aware of other results analogous to Theorem 2.7 in the literature.

As noted in the text, the CES production function was first introduced by Arrow et al. (1961). This production function plays an important role in many applied macroeconomic and economic growth models. The Inada conditions introduced in Assumption 2 are from Inada (1963).

Finally, the interested reader should look at the paper by Hakenes and Irmen (2006) for why Inada conditions can introduce an additional equilibrium path (other than the no-activity equilibrium) at $k = 0$ in continuous time even when $f(0) = 0$. Here it suffices to say that whether this steady state exists is a matter of the order in which limits are taken. In any case, as noted in the text, the steady state at $k = 0$ has no economic content and is ignored throughout the book.

2.11 Exercises

2.1 Show that competitive labor markets and Assumption 1 imply that the wage rate must be strictly positive and thus (2.4) implies (2.3).

2.2 Prove that Assumption 1 implies that $F(A, K, L)$ is concave in K and L but not strictly so.

2.3 Show that when F exhibits constant returns to scale and factor markets are competitive, the maximization problem in (2.5) either has no solution (the firm can make infinite profits), a unique solution $K = L = 0$, or a continuum of solutions (i.e., any (K, L) with $K/L = \kappa$ for some $\kappa > 0$ is a solution).

2.4 Consider the Solow growth model in continuous time with the following per capita production function:

$$f(k) = k^4 - 6k^3 + 11k^2 - 6k.$$

(a) Which parts of Assumptions 1 and 2 does the underlying production function $F(K, L)$ violate?

(b) Show that with this production function, there exist three steady-state equilibria.

(c) Prove that two of these steady-state equilibria are locally stable, while one of them is locally unstable. Can any of these steady-state equilibria be globally stable?

2.5 Prove Proposition 2.7.

2.6 Prove Proposition 2.8.

2.7 Let us introduce government spending in the basic Solow model. Consider the basic model without technological change and suppose that (2.9) takes the form

$$Y(t) = C(t) + I(t) + G(t),$$

with $G(t)$ denoting government spending at time t. Imagine that government spending is given by $G(t) = \sigma Y(t)$.

(a) Discuss how the relationship between income and consumption should be changed. Is it reasonable to assume that $C(t) = sY(t)$?

(b) Suppose that government spending partly comes out of private consumption, so that $C(t) = (s - \lambda\sigma)Y(t)$, where $\lambda \in [0, 1]$. What is the effect of higher government spending (in the form of higher σ) on the equilibrium of the Solow model?

(c) Now suppose that a fraction ϕ of $G(t)$ is invested in the capital stock, so that total investment at time t is given by

$$I(t) = (1 - s - (1 - \lambda)\,\sigma + \phi\sigma)\,Y(t).$$

Show that if ϕ is sufficiently high, the steady-state level of capital-labor ratio will increase as a result of higher government spending (corresponding to higher σ). Is this reasonable? How would you alternatively introduce public investments in this model?

2.8 Suppose that $F(K, L, A)$ is concave in K and L (though not necessarily strictly so) and satisfies Assumption 2. Prove Propositions 2.2 and 2.5. How do we need to modify Proposition 2.6?

2.9 Prove Proposition 2.6.

2.10 Prove Corollary 2.2.

2.11 Consider a modified version of the continuous-time Solow growth model where the aggregate production function is

$$F(K, L, Z) = L^{\beta} K^{\alpha} Z^{1-\alpha-\beta},$$

where Z is land, available in fixed inelastic supply. Assume that $\alpha + \beta < 1$, capital depreciates at the rate δ, and there is an exogenous saving rate of s.

(a) First suppose that there is no population growth. Find the steady-state capital-labor ratio in the steady-state output level. Prove that the steady state is unique and globally stable.

(b) Now suppose that there is population growth at the rate n, that is, $\dot{L}/L = n$. What happens to the capital-labor ratio and output level as $t \to \infty$? What happens to returns to land and the wage rate as $t \to \infty$?

(c) Would you expect the population growth rate n or the saving rate s to change over time in this economy? If so, how?

2.12 Consider the continuous-time Solow model without technological progress and with constant rate of population growth equal to n. Suppose that the production function satisfies Assumptions 1 and 2. Assume that capital is owned by capitalists and labor is supplied by a different set of agents, the workers. Following a suggestion by Kaldor (1957), suppose that capitalists save a fraction s_K of their income, while workers consume all of their income.

(a) Define and characterize the steady-state equilibrium of this economy and study its stability.

(b) What is the relationship between the steady-state capital-labor ratio k^* and the golden rule capital stock k^*_{gold} defined in Section 2.2.3?

2.13 Let us now make the opposite assumption of Exercise 2.12 and suppose that there is a constant saving rate $s \in (0, 1)$ out of labor income and no savings out of capital income. Suppose that the aggregate production function satisfies Assumptions 1 and 2. Show that in this case multiple steady-state equilibria are possible.

* 2.14 In this exercise, you are asked to generalize Theorem 2.6 to a situation in which, rather than

$$\dot{Y}(t)/Y(t) = g_Y > 0,\ \dot{K}(t)/K(t) = g_K > 0, \text{ and } \dot{C}(t)/C(t) = g_C > 0$$

for all $t \geq T$ with $T < \infty$, we have

$$\dot{Y}(t)/Y(t) \to g_Y > 0,\ \dot{K}(t)/K(t) \to g_K > 0, \text{ and} \dot{C}(t)/C(t) \to g_C > 0.$$

(a) Show, by constructing a counterexample, that Part 1 of Theorem 2.6 is no longer correct without further conditions. [Hint: consider $g_c < g_k = g_y$.] What conditions need to be imposed to ensure that these limiting growth rates are equal to one another?

(b) Now suppose that Part 1 of Theorem 2.6 has been established (in particular, $g_Y = g_K$). Show that the equivalent of the steps in the proof of the theorem imply that for any T and $t \geq T$, we have

$$\exp\left(-\int_T^t g_Y(s)ds\right) Y(t)$$
$$= \tilde{F}\left[\exp\left(-\int_T^t g_K(s)ds\right) K(t), \exp\left(-n(t-T)\right) L(t), \tilde{A}(T)\right],$$

where $g_Y(t) \equiv \dot{Y}(t)/Y(t)$, and $g_K(t)$ and $g_C(t)$ are defined similarly. Then show that

$$Y(t) = \tilde{F}\left[\exp\left(\int_T^t \left(g_Y(s) - g_K(s)\right) ds\right) K(t), \exp\left(\int_T^t \left(g_Y(s) - n\right) ds\right) L(t), \tilde{A}(T)\right].$$

Next observe that for any $\varepsilon_T > 0$, there exists $T < \infty$, such that $\left|g_Y(t) - g_Y\right| < \varepsilon_T/2$ and $\left|g_K(t) - g_Y\right| < \varepsilon_T/2$ (from the hypotheses that $\dot{Y}(t)/Y(t) \to g_Y > 0$ and $\dot{K}(t)/K(t) \to g_K > 0$). Consider a sequence (or net; see Appendix A) $\left\{\varepsilon_T\right\} \to 0$, which naturally corresponds to $T \to \infty$ in the above definition. Take $t = \xi T$ for some $\xi > 1$, and show that Part 2 of Theorem 2.6 holds if $\varepsilon_T T \to 0$ (as $T \to \infty$). Using this argument, show that if both $g_Y(t)$ and $g_K(t)$ converge to g_Y and g_K at a rate strictly faster than $1/t$, the asymptotic production function has a representation of the form $F(K(t), A(t)L(t))$, but that this conclusion does not hold if either $g_Y(t)$ or $g_K(t)$ converges at a slower rate. [Hint: here an asymptotic representation means that $\lim_{t\to\infty} \tilde{F}/F = 1$.]

2.15 Recall the definition of the elasticity of substitution σ in (2.37). Suppose labor markets are competitive and the wage rate is equal to w. Prove that if the aggregate production function $F(K, L, A)$ exhibits constant returns to scale in K and L, then

$$\varepsilon_{y,w} \equiv \frac{\partial y/\partial w}{y/w} = \sigma,$$

where, as usual, $y \equiv F(K, L, A)/L$.

* 2.16 In this exercise you are asked to derive the CES production function (2.38) following the method in the original article by Arrow et al. (1961). These authors noted that a good empirical approximation to the relationship between income per capita and the wage rate was provided by an equation of the form

$$y = \alpha w^{\sigma},$$

where $y = f(k)$ is again output per capita and w is the wage rate. With competitive markets, recall that $w = f(k) - kf'(k)$. Thus the above equation can be written as

$$y = \alpha(y - ky')^{\sigma},$$

where $y = y(k) \equiv f(k)$ and y' denotes $f'(k)$. This is a nonlinear first-order differential equation.

(a) Using separation of variables (see Appendix B), show that the solution to this equation satisfies

$$y(k) = \left(\alpha^{-1/\sigma} + c_0 k^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},$$

where c_0 is a constant of integration.

(b) Explain how you would put more structure on α and c_0 and derive the exact form of the CES production function in (2.38).

2.17 Consider the Solow growth model with constant saving rate s and depreciation rate of capital equal to δ. Assume that population is constant and the aggregate production function is given by the constant returns to scale production function

$$F(A_K(t)K(t), A_L(t)L(t)),$$

where $\dot{A}_L(t)/A_L(t) = g_L > 0$ and $\dot{A}_K(t)/A_K(t) = g_K > 0$.

(a) Suppose that F is Cobb-Douglas. Determine the BGP growth rate and the adjustment of the economy to the steady state.

(b) Suppose that F is not Cobb-Douglas (even asymptotically). Prove that there cannot exist $T < \infty$ such that the economy is on a BGP for all $t \geq T$. Explain why.

* 2.18 Consider the environment in Exercise 2.17. Suppose that F takes a CES form as in (2.38) with the elasticity of substitution between capital and labor $\sigma < 1$, $g_K > g_L$, and there is a constant saving rate s. Show that as $t \to \infty$, the economy converges to a BGP where the share of labor in national income is equal to 1 and capital, output, and consumption all grow at the rate g_L. In light of this result, discuss the claim in the literature that capital-augmenting technological change is inconsistent with balanced growth. Why is the claim in the literature incorrect? Relate your answer to Exercise 2.14.

* 2.19 In the context of Theorem 2.6, consider the production function

$$\tilde{F}(K(t), L(t), \tilde{A}(t)) = K(t)^{\tilde{A}(t)} L(t)^{1-\tilde{A}(t)},$$

where $\tilde{A}(t) : \mathbb{R}_+ \to (0, 1)$ is an arbitrary function of time, representing technology.

(a) Show that when $K(t) = \exp(nt)$ and $L(t) = \exp(nt)$ for $n \geq 0$, the conditions of Theorem 2.6 are satisfied and $\tilde{F}$ has a representation of the form $F(K(t), A(t)L(t))$. Determine a class of functions that provide such a representation.

(b) Show that the derivatives of $\tilde{F}$ and F are not equal.

(c) Suppose that factor markets are competitive. Show that while capital, output, and consumption grow at a constant rate, the capital share in national income behaves in an arbitrary fashion. [Hint: consider, for example, $\tilde{A}(t) = (2 + \sin(t))/4$.]

2.20 Consider the Solow model with noncompetitive labor markets. In particular, suppose that there is no population growth and no technological progress and output is given by $F(K, L)$. The saving rate is equal to s and the depreciation rate is given by δ.

(a) First suppose that there is a minimum wage $\bar{w}$, such that workers are not allowed to be paid less than $\bar{w}$. If labor demand at this wage falls short of L, employment is equal to the amount of labor demanded by firms, L^d (and the unemployed do not contribute to production and earn zero). Assume that $\bar{w} > f(k^*) - k^* f'(k^*)$, where k^* is the steady-state capital-labor ratio of the basic Solow model given by $f(k^*)/k^* = \delta/s$. Characterize the dynamic equilibrium path of this economy starting with some amount of physical capital $K(0) > 0$.

(b) Next consider a different form of labor market imperfection, whereby workers receive a fraction $\lambda > 0$ of output of their employer as their wage income. Characterize a dynamic equilibrium path in this case. [Hint: recall that the saving rate is still equal to s.]

2.21 Consider the discrete-time Solow growth model with constant population growth at the rate n, no technological change, and full depreciation (i.e., $\delta = 1$). Assume that the saving rate is a function of the capital-labor ratio and is thus given by $s(k)$.

(a) Suppose that $f(k) = Ak$ and $s(k) = s_0 k^{-1} - 1$. Show that if $A + \delta - n = 2$, then for any $k(0) \in (0, As_0/(1+n))$, the economy immediately settles into an asymptotic cycle and continuously fluctuates between $k(0)$ and $As_0/(1+n) - k(0)$. (Suppose that $k(0)$ and the parameters are given such that $s(k) \in (0, 1)$ for both $k = k(0)$ and $k = As_0/(1+n) - k(0)$.)

(b) Now consider the more general continuous production function $f(k)$ and saving function $s(k)$, such that there exist $k_1, k_2 \in \mathbb{R}_+$ with $k_1 \neq k_2$ and

$$k_2 = \frac{s(k_1) f(k_1) + (1-\delta) k_1}{1+n},$$

$$k_1 = \frac{s(k_2) f(k_2) + (1-\delta) k_2}{1+n}.$$

Show that when such (k_1, k_2) exist, there may also exist a stable steady state.

(c) Prove that such cycles are not possible in the continuous-time Solow growth model for any (possibly non-neoclassical) continuous production function $f(k)$ and continuous $s(k)$. [Hint: consider the equivalent of Figure 2.9.]

(d) What does the result in parts a–c imply for the approximations of discrete time by continuous time suggested in Section 2.4?

(e) In light of your answer to part d, what do you think of the cycles in parts a and b?

(f) Show that if $f(k)$ is nondecreasing in k and $s(k) = k$, cycles as in parts a and b are not possible in discrete time either.

2.22 Characterize the asymptotic equilibrium of the modified Solow/AK model mentioned in Section 2.6, with a constant saving rate s, depreciation rate δ, no population growth, and an aggregate production function of the form

$$F(K(t), L(t)) = A_K K(t) + A_L L(t).$$

2.23 Consider the basic Solow growth model with a constant saving rate s, constant population growth at the rate n, and no technological change, and suppose that the aggregate production function takes the CES form in (2.38).

(a) Suppose that $\sigma > 1$. Show that in this case equilibrium behavior can be similar to that in Exercise 2.22 with sustained growth in the long run. Interpret this result.

(b) Now suppose that $\sigma \to 0$, so that the production function becomes Leontief:

$$Y(t) = \min\left\{\gamma A_K(t) K(t); (1 - \gamma)\ A_L(t) L(t)\right\}.$$

The model is then identical to the classical Harrod-Domar growth model developed by Roy Harrod and Evsey Domar (Harrod, 1939; Domar, 1946). Show that in this case there is typically no steady-state equilibrium with full employment and no idle capital. What happens to factor prices in these cases? Explain why this case is pathological, giving at least two reasons for why we may expect equilibria with idle capital or idle labor not to apply in practice.

2.24 Show that the CES production function (2.38) violates Assumption 2 unless $\sigma = 1$.

2.25 Prove Proposition 2.12.

2.26 Prove Proposition 2.13.

2.27 In this exercise, we work through an alternative conception of technology, which will be useful in the next chapter. Consider the basic Solow model in continuous time and suppose that $A(t) = A$, so that there is no technological progress of the usual kind. However, assume that the relationship between investment and capital accumulation is modified to

$$\dot{K}(t) = q(t) I(t) - \delta K(t),$$

where $\left[q(t)\right]_{t=0}^{\infty}$ is an exogenously given time-varying path (function). Intuitively, when $q(t)$ is high, the same investment expenditure translates into a greater increase in the capital stock.

Therefore we can think of $q(t)$ as the inverse of the relative price of machinery to output. When $q(t)$ is high, machinery is relatively cheaper. Gordon (1990) documented that the relative prices of durable machinery have been declining relative to output throughout the postwar era. This decline is quite plausible, especially given recent experience with the decline in the relative price of computer hardware and software. Thus we may want to suppose that $\dot{q}(t) > 0$. This exercise asks you to work through a model with this feature based on Greenwood, Hercowitz, and Krusell (1997).

(a) Suppose that $\dot{q}(t)/q(t) = \gamma_K > 0$. Show that for a general production function, $F(K, L)$, there exists no BGP.

(b) Now suppose that the production function is Cobb-Douglas, $F(K, L) = K^{\alpha}L^{1-\alpha}$, and characterize the unique BGP.

(c) Show that this steady-state equilibrium does not satisfy the Kaldor fact of constant K/Y. Is this discrepancy a problem? [Hint: how is K measured in practice? How is it measured in this model?]